

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:14:34 ; Search time 19.84 Seconds
(without alignments)
92,401 Million cell updates/sec

Title: US-09-822-110-5

Perfect score: 253
Sequence: 1 PFIRNAKPVSLRLITLTEAM.....QRELEEEBNDDELDLSHT 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	253	100.0	491	1 STK3_HUMAN	Q13188 homo sapien
2	165.5	65.4	487	1 STK4_HUMAN	Q13043 homo sapien
3	71.5	28.3	312	1 PHO4_YEAST	P07270 saccharomyc
4	70	27.7	626	1 ALI2_ARABY	P43238 arachis hyp
5	69	27.3	200	1 HMG1_CHICK	P36194 gallus gall
6	69	27.3	201	1 HMG2_CHICK	P40618 gallus gall
7	69	27.3	612	1 RPSD_BUCAL	P57163 buchera ap
8	68.5	26.9	324	1 SIAL_MOUSE	Q61711 mus musculu
9	68	26.7	678	1 GARP_PLAIF	P13816 plasmodium
10	67.5	26.7	677	1 OS9_HUMAN	Q13438 homo sapien
11	67.5	26.7	879	1 MCN3_SCHPO	P30666 schizosacch
12	67	26.5	178	1 NPM3_HUMAN	Q75607 homo sapien
13	67	26.5	416	1 CRRC_RAT	P18418 rattus norv
14	67	26.5	418	1 YSE2_CABEL	Q09936 caenorhabdi
15	66.5	26.3	879	1 SPI0_HUMAN	P23437 homo sapien
16	66.5	26.3	890	1 IF2_ECOLI	P02995 escherichia
17	66	26.1	598	1 ABB1_RAT	P46933 rattus norv
18	66	26.1	706	1 NUCL_MOUSE	P09405 mus musculu
19	66	26.1	1121	1 MYT1_HUMAN	Q01585 homo sapien
20	65	25.7	182	1 YADA_SCHPO	Q09838 schizosacch
21	65	25.7	310	1 SIAL_BOVIN	Q28862 bos taurus
22	65	25.7	976	1 VPA1_ARATH	P33043 arabidopsis
23	64.5	25.5	721	1 YCF2_OENPI	P31568 oenothera p
24	64	25.3	180	1 HMG1_CRIGR	P07156 cricetus
25	64	25.3	209	1 HMG2_PIG	P17741 sus scrofa
26	64	25.3	214	1 HMG1_MOUSE	P07155 mus musculu
27	64	25.3	320	1 SIAL_RAT	P13819 rattus norv
28	64	25.3	416	1 CRRC_MOUSE	P14211 mus musculu
29	64	25.3	712	1 GLCX_SOYBN	P11827 glycine max
30	64	25.3	712	1 NUCL_RAT	P13383 rattus norv
31	63.5	25.1	1161	1 NRDC_RAT	P47245 rattus norv
32	63	24.9	240	1 SNF7_YEAST	P39929 saccharomyc
33	63	24.9	299	1 NPM_XENLA	P07222 xenopus lae

ALIGNMENTS

RESULT	1	STANDARD	PRT	491 AA
ID	STK3_HUMAN	Q13188; Q15801; Q15445;		
AC	Q13188; Q15801; Q15445;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DE	SERINE/THREONINE PROTEIN KINASE 3 (EC 2.7.1.37) (STE20-LIKE KINASE			
DE	MST2) (MST-2) (MAMMALIAN STE20-LIKE PROTEIN KINASE 2)			
DE	(SERINE/THREONINE PROTEIN KINASE KRS-1).			
GN	STK3 OR MST2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96144292; PubMed=8566796;			
RA	Creasy C.L., Chernoff J.,			
RT	"Cloning and characterization of a member of the MST subfamily of			
RT	Ste20-like kinases."			
RT	Gene=167:303-306(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96413604; PubMed=8816758;			
RA	Taylor L.K., Wang H.C., Erikson R.L.,			
RT	"Newly identified stress-responsive protein kinases, Krs-1 and Krs-			
RT	2."			
RT	Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996)			
RN	[3]			
RP	SEQUENCE OF 96-203 FROM N.A.			
RX	MEDLINE=94100173; PubMed=8274451;			
RA	Schultz S.J., Nigg E.A.,			
RT	"Identification of 21 novel human protein kinases, including 3 members			
RT	of a family related to the cell cycle regulator nima of Aspergillus			
RT	nidulans."			
RT	Cell Growth Differ. 4:821-830(1993).			
CC	- FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT			
CC	MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY			
CC	SIMILARITY).			
CC	- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.			
CC	- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).			
CC	- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,			
CC	SKELTAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT			
CC	HEART, LUNG AND BRAIN TISSUES.			
CC	- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	STE20 SUBFAMILY.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isdb.ch/announce/			
CC	or send an email to license@isdb.ch).			
CC				

DR EMBL: U26424; AAC50386.1; -
 DR EMBL: U60206; AAB17261.1; -
 DR EMBL: Z25422; CA80909.1; -
 DR HSSP: P00518; IPRK.
 DR MIM: 605030; -
 DR InterPro: IPR000719; Euk PKinase.
 DR InterPro: IPR002290; Ser Thr Kin_actsite.
 DR Pfam: PF00069; pkinase.1
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase: Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 27 278
 FT NP_BIND 33 41
 FT BINDING 56 56
 FT ACT_SITE 146 146
 FT DOMAIN 308 314
 FT DOMAIN 370 375
 FT CONFLICT 96 98
 FT CONFLICT 121 121
 FT CONFLICT 203 203
 FT CONFLICT 303 303
 FT CONFLICT 332 334
 FT SEQUENCE 491 AA; 56261 MM; 9CA3B0644F3C14A9 CRC64;

Query Match 100.0%; Score:253; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 4.3e-17;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PFKNKAPVSIILDLITFAMEIKAKRHDEQORLEEEENSDELDSDST 50
 DB 276 PFKNKAPVSIILDLITFAMEIKAKRHDEQORLEEEENSDELDSDST 325

RESULT 2

STK4_HUMAN STANDARD; PRT; 487 AA.

AC Q13043; Q15802; Q9NT24;
 DT 20-AUG-2001 (Rel. 40; Created)
 DT 20-AUG-2001 (Rel. 40; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE SERINE/THREONINE PROTEIN KINASE 4 (EC 2.7.1.37) (STE20-LIKE KINASE
 DE MST1) (MST1) (MAMMALIAN STE20-LIKE PROTEIN KINASE 1)
 DE (SERINE/THREONINE PROTEIN KINASE KRS-2).
 DE STK4 OR MST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394929; PubMed=7665586;
 RA Creasy C.L.; Chernoff J.;
 RT "Cloning and characterization of a human protein kinase with homology
 RT to Ste20."
 RL J. Biol. Chem. 270:21695-21700(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413604; PubMed=8816758;
 RA Taylor L.K.; Wang H.C.; Erikson R.L.;
 RT "Newly identified stress-responsive protein kinases: Krs-1 and Krs-
 RT 2."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
 RN [3]
 RP SEQUENCE OF 1-435 FROM N.A.
 RA Laird G.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.
 CC -1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
 CC THE KINASE ACTIVITY.

CC -1- SURCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED.
 CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC STE20 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: U18297; AAB83254.1; -
 DR EMBL: U60207; AAB17262.1; -
 DR EMBL: AL109839; CAB89421.1; -
 DR HSSP: P24941; IHCL.
 DR MIM: 604965; -
 DR InterPro: IPR000719; Euk PKinase.
 DR InterPro: IPR002290; Ser Thr Kin_actsite.
 DR Pfam: PF00069; pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transferrase: Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 30 281
 FT NP_BIND 36 44
 FT BINDING 59 59
 FT ACT_SITE 149 149
 FT DOMAIN 373 378
 FT CONFLICT 222 222
 FT CONFLICT 312 312
 FT SEQUENCE 487 AA; 55630 MM; 150758BEC5F77D5C CRC64;

Query Match 65.4%; Score 165.5; DB 1; Length 487;
 Best Local Similarity 62.7%; Pred. No. 7.5e-09;
 Matches 32; Conservative 12; Mismatches 6; Indels 1; Gaps 1;

OY 1 PFKNKAPVSIILDLITFAMEIKAKRHDEQORLEEEENSDELDSDST 50
 DB 279 PFVRSKAGVSIILDLITFAMEIKAKRHDEQORLEEEENSDELDSDST 329

RESULT 3

PHO4_YEAST STANDARD; PRT; 312 AA.

AC P07270;
 DT 01-APR-1988 (Rel. 07; Created)
 DT 01-FEB-1991 (Rel. 17; Last sequence update)
 DT 15-DEC-1998 (Rel. 37; Last annotation update)
 DE PHOSPHATE SYSTEM POSITIVE REGULATORY PROTEIN PHO4.
 DE PHO4 OR YPR034C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 OC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86176785; PubMed=3008105;
 RA Legrain M.; de Wilde M.; Hilger F.;
 RT "Isolation, physical characterization and expression analysis of the
 RT Saccharomyces cerevisiae positive regulatory gene PHO4."
 RL Nucleic Acids Res. 14:3059-3073(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89364688; PubMed=2505053;
 RA Yoshida K.; Kuromitsu Z.; Ogawa N.; Oshima Y.;
 RT "Mode of expression of the positive regulatory genes PHO2 and PHO4 of

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:04:59 ; Search time 26.93 Seconds
(without alignments)
41.781 Million cell updates/sec

Title: US-09-822-110-5

Perfect score: 253
Sequence: 1 PFIRAKKPVSLRLDILTEAM.....ORELEEEENSDELDLSHT 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166.5	65.8	487	2	US-08-712-709-8
2	166.5	65.8	487	3	US-09-111-444-8
3	166.5	65.8	487	4	US-09-541-228-8
4	67	26.5	416	2	US-09-211-930-11
5	67	26.5	416	3	US-09-340-993-11
6	67	26.5	416	4	US-09-468-442-11
7	67	26.5	1162	2	US-08-728-333A-2
8	65	25.7	416	2	US-09-211-930-3
9	65	25.7	416	3	US-09-340-993-3
10	65	25.7	416	4	US-09-468-442-3
11	65	25.7	759	1	US-08-676-967-1
12	65	25.7	759	1	US-08-676-974-1
13	65	25.7	759	2	US-09-098-487-1
14	64	25.3	1719	2	US-08-459-568-4
15	64	25.3	1719	2	US-08-399-411-4
16	64	25.3	1719	3	US-08-516-859A-4
17	62.5	24.7	379	1	US-08-552-142A-11
18	62.5	24.7	381	1	US-08-910-973-11
19	62.5	24.7	382	3	US-09-224-332-9
20	62	24.5	353	3	US-08-688-988-31
21	62	24.5	781	2	US-08-373-134D-2
22	62	24.5	781	2	US-09-114-637-2
23	62	24.5	3135	1	US-08-323-170B-2
24	61	24.1	270	2	US-08-853-743-5
25	61	24.1	270	2	US-09-185-370-5
26	61	24.1	285	4	US-09-282-305-18
27	61	24.1	905	2	US-08-574-959A-9

28	61	24.1	905	4	US-09-357-014-9	Sequence 9, Appli
29	61	24.1	1135	2	US-08-574-959A-7	Sequence 7, Appli
30	61	24.1	1135	4	US-09-357-014-7	Sequence 9, Appli
31	60.5	23.9	675	1	US-08-317-522A-9	Sequence 9, Appli
32	60.5	23.9	675	1	US-08-439-818A-9	Sequence 9, Appli
33	60.5	23.9	675	2	US-08-751-965-9	Sequence 9, Appli
34	60.5	23.9	675	2	US-08-738-975-9	Sequence 9, Appli
35	60.5	23.9	675	2	US-08-728-626-9	Sequence 9, Appli
36	60.5	23.9	675	3	US-08-808-599A-9	Sequence 9, Appli
37	60	23.7	240	3	US-08-114-555A-8	Sequence 8, Appli
38	60	23.7	240	3	US-08-559-397A-14	Sequence 14, Appli
39	60	23.7	303	4	US-09-045-973-1	Sequence 1, Appli
40	60	23.7	522	4	US-09-075-272-3	Sequence 3, Appli
41	60	23.7	622	4	US-09-075-272-2	Sequence 2, Appli
42	60	23.7	887	1	US-07-867-106-3	Sequence 3, Appli
43	60	23.7	1085	2	US-08-431-080-28	Sequence 28, Appli
44	60	23.7	1085	2	US-08-938-534-28	Sequence 28, Appli
45	59.5	23.5	310	2	US-08-943-600A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-712-709-8
Sequence 8, Application US/08712709
Patent No. 5863780
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,709
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791
US-08-712-709-8

Query Match 65.8% Score 166.5; DB 2; Length 487;
Best local similarity 62.7%; Pred. No. 2.5e12;
Matches 32; Conservative 12; Mismatches 6; Indels 1; Gaps 1;
Cy 1 PFIRAKKPVSLRLDILTEAMIKAKRHDEQREL-EEEEENSDELDLSHT 50

Db 279 PFVRSAGVSIIRDLINEMAMVKLRQESQOREMDQDEENSEDEMDSGT 329

RESULT 2

US-09-111-444-8

Sequence 8, Application US/09111444
Patent No. 6045792

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: IBM Compatible

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/111,444

FILING DATE:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/712,709

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 487 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1117791

US-09-111-444-8

Query Match 65.8%; Score 166.5; DB 3; Length 487;
Best Local Similarity 62.7%; Pred. No. 2.5e-12;
Matches 32; Conservative 12; Mismatches 6; Indels 1; Gaps 1;

OY 1 PFIRAKPVSIIRDLITEAMEIKARRHDEQOREL-EEEEENSEDEDELDSHT 50
Db 279 PFVRSAGVSIIRDLINEMAMVKLRQESQOREMDQDEENSEDEMDSGT 329

RESULT 3

US-09-541-228-8

Sequence 8, Application US/09541228
Patent No. 6232077

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/541,228

FILING DATE:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/712,709

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 487 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1117791

US-09-541-228-8

Query Match 65.8%; Score 166.5; DB 4; Length 487;
Best Local Similarity 62.7%; Pred. No. 2.5e-12;
Matches 32; Conservative 12; Mismatches 6; Indels 1; Gaps 1;

OY 1 PFIRAKPVSIIRDLITEAMEIKARRHDEQOREL-EEEEENSEDEDELDSHT 50
Db 279 PFVRSAGVSIIRDLINEMAMVKLRQESQOREMDQDEENSEDEMDSGT 329

RESULT 4

US-09-211-930-11

Sequence 11, Application US/09211930
Patent No. 5962265

GENERAL INFORMATION:

APPLICANT: Tyrell E. No. 5962265r1s

APPLICANT: William Craig Moore

TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION

FILE REFERENCE: PHM 70296

CURRENT APPLICATION NUMBER: US/09/211,930

CURRENT FILING DATE: 1998-12-15

EARLIER APPLICATION NUMBER: GB 9726851.0

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 11

LENGTH: 416

TYPE: PRT

ORGANISM: Mus musculus

US-09-211-930-11

Query Match 26.5%; Score 67; DB 2; Length 416;
Best Local Similarity 29.8%; Pred. No. 1.4;
Matches 14; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:06:00 ; Search time 54.15 Seconds

(without alignments)
68.396 Million cell updates/sec

Title: US-09-822-110-5

Perfect score: 253
Sequence: 1 PPIKAKPVSYILRLITEAM.....ORELEEEENSDDELDSHT 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSB8/gcgdata/geneseq/AA1980.DAT:*
2: /SIDSB8/gcgdata/geneseq/AA1981.DAT:*
3: /SIDSB8/gcgdata/geneseq/AA1982.DAT:*
4: /SIDSB8/gcgdata/geneseq/AA1983.DAT:*
5: /SIDSB8/gcgdata/geneseq/AA1984.DAT:*
6: /SIDSB8/gcgdata/geneseq/AA1985.DAT:*
7: /SIDSB8/gcgdata/geneseq/AA1986.DAT:*
8: /SIDSB8/gcgdata/geneseq/AA1987.DAT:*
9: /SIDSB8/gcgdata/geneseq/AA1988.DAT:*
10: /SIDSB8/gcgdata/geneseq/AA1989.DAT:*
11: /SIDSB8/gcgdata/geneseq/AA1990.DAT:*
12: /SIDSB8/gcgdata/geneseq/AA1991.DAT:*
13: /SIDSB8/gcgdata/geneseq/AA1992.DAT:*
14: /SIDSB8/gcgdata/geneseq/AA1993.DAT:*
15: /SIDSB8/gcgdata/geneseq/AA1994.DAT:*
16: /SIDSB8/gcgdata/geneseq/AA1995.DAT:*
17: /SIDSB8/gcgdata/geneseq/AA1996.DAT:*
18: /SIDSB8/gcgdata/geneseq/AA1997.DAT:*
19: /SIDSB8/gcgdata/geneseq/AA1998.DAT:*
20: /SIDSB8/gcgdata/geneseq/AA1999.DAT:*
21: /SIDSB8/gcgdata/geneseq/AA2000.DAT:*
22: /SIDSB8/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	167.5	66.2	487 20	AAV21674 Human Ste20 homolo
2	74	29.2	524 18	AAW32096 Miniature swine re
3	74	29.2	524 22	AAW3285 Retroviral protein
4	71	28.1	142 21	AAW41902 Arabidopsis thalia
5	71	28.1	261 21	AAW41901 Arabidopsis thalia
6	71	28.1	280 21	AAW41900 Arabidopsis thalia
7	70	27.7	415 20	AAW40913 Ara h 1 allergen p
8	70	27.7	626 18	AAW2150 Peanut allergen Ar
9	70	27.7	626 20	AAW15244 Peanut allergen, A
10	70	27.7	626 20	AAW25657 Peanut allergen, I
11	70	27.7	626 22	AAW04706 Anaphylactic anti

12	70	27.7	634	21	AAB33599	Modified Ara h 1 a
13	70	27.7	634	22	AAU04709	Modified anaphylac
14	70	27.7	634	22	AAU05034	Modified anaphylac
15	69	27.3	117	22	AAW36518	Peptide #10555 enc
16	69	27.3	524	18	AAW39271	Porcine retrovirus
17	69	27.3	524	18	AAW32091	Porcine retrovirus
18	69	27.3	524	18	AAW32091	Porcine retrovirus
19	69	27.3	524	22	AAW70629	Porcine endogenous
20	67	26.5	162	21	AAW33282	Defective retrovir
21	67	26.5	162	21	AAW33282	Defective retrovir
22	67	26.5	214	21	AAW21047	Human nucleic acid
23	67	26.5	241	21	AAW39981	Arabidopsis thalia
24	67	26.5	247	21	AAW29980	Arabidopsis thalia
25	67	26.5	416	20	AAW22651	A murine signal tr
26	67	26.5	416	21	AAW21673	Murine 5e.new poly
27	67	26.5	416	21	AAW82277	Mouse protein sequ
28	67	26.5	1162	21	AAW6255	Kaposi's sarcoma-a
29	67	26.5	1162	21	AAW58500	HHV8 ORF 73 protel
30	66.5	26.3	993	22	AAW62331	Amino acid sequenc
31	66.5	26.3	1312	22	AAW59437	Human polypeptide
32	66.5	26.3	1312	22	AAW76884	Human polypeptide
33	66.5	26.3	1312	22	AAW38967	Human polypeptide
34	66.5	26.1	219	21	AAW33517	Arabidopsis thalia
35	66.5	26.1	219	21	AAW33516	Arabidopsis thalia
36	66.5	26.1	309	21	AAW33515	Arabidopsis thalia
37	66.5	26.1	309	21	AAW70491	Leucocytoczoan prot
38	65.5	25.9	1151	20	AAW5039	Human N-arginine d
39	65.5	25.9	329	21	AAW37724	Arabidopsis thalia
40	65.5	25.9	360	21	AAW37722	Arabidopsis thalia
41	65.5	25.7	65	22	AAW15961	Peptide #2395 enco
42	65.5	25.7	65	22	AAW28463	Peptide #2300 enco
43	65.5	25.7	65	22	AAW03698	Peptide #2380 enco
44	65.5	25.7	228	22	AAW63361	Human breast cance
45	65.5	25.7	231	22	AAW94376	Human protein sequ

ALIGNMENTS

RESULT 1	AAV21674	AAV21674 standard; Protein: 487 AA.
ID	AAV21674	
AC	AAV21674	
XX		
XX		18-AUG-1999 (first entry)
DE		Human Ste20 homologue polypeptide S201.
XX		
KW		Sterile 20; serine/threonine kinase; C12.2bs; 5e.new; murine;
KW		erythroleukemia; mitogenic signalling; cell differentiation; metastasis;
KW		signaling pathway; cancer; ischemic stroke; heart disease; inflammation.
XX		
OS		Homo sapiens.
XX		
PN		W09929857-A1.
XX		
PD		17-JUN-1999.
XX		
PF		09-DEC-1998; 98W0-US26116.
XX		
PR		09-DEC-1997; 97US-0069078.
XX		
PA		(CHIL-) CHILDRENS MEDICAL CENT.
XX		
PI		Agarwal S, Best J, Vail B, Zon LI;
XX		
DR		MP1: 1999-385605/32.
XX		
PT		Murine homologues of the Sterile 20 family of serine/threonine
PT		kinases, useful in elucidation of mitogenic signalling pathways
XX		
PS		Disclosure: Fig 4; 37pp; English.

xx The invention provides two members of the sterile 20 (Ste20) family of
 CC serine/threonine kinases (C12.2bs and 5e.new). Host cells transfected
 CC with vectors comprising the nucleic acids encoding the polypeptides are
 CC used for their recombinant expression. The Ste20 kinases, isolated from a
 CC murine erythroleukemia cDNA library, are regulatory molecules involved in
 CC mitogenic signaling as well as other cellular phenomena such as
 CC morphology and motility. These activities are important factors in
 CC development, cell differentiation, cancer and metastases. Manipulation of
 CC the C12.2bs and 5e.new proteins or polynucleotides are useful for
 CC manipulation of the signaling pathways involved and will allow
 CC development of reagents to modulate these signaling pathways. The kinases
 CC are also important for ischemic stroke, heart disease and inflammation.
 CC The antibodies can be used to inhibit the activity of the kinase,
 CC especially in vitro and in cell extracts, as well as identifying the kinases
 CC in immunoassays. The present sequence represents a S201 polypeptide, a
 CC human homologue of Ste20.

xx
 SQ Sequence 487 AA;

Query Match 66.2%; Score 167.5; DB 20; Length 487;
 Best Local Similarity 64.7%; Pred. No. 5.4e-12;
 Matches 33; Conservative 11; Mismatches 6; Indels 1; Gaps 1;

OY 1 PFIKNAKPVSLDLITEAMEIKAKRHDEQOREL-EEEEENSDDEDLSHT 50
 DB 279 ptkvksakysilrldlineamdvklkrqeqvqdqddenseedamdsgt 329

RESULT 2

AAW32096 AAW32096 standard; Protein; 524 AA.

AC AAW32096;

DT 09-FEB-1998 (first entry)

DE Miniature swine retrovirus GAG;protein.

XX Retrovirus; porcine; GAG protein; xenotransplantation; infectious;

KW provirus; organ transplant; donor; activated virus; PCR.

OS Porcine retrovirus.

XX Key Location/Qualifiers

FT Protein 1.524

XX /Label= GAG_protein

PN WO9721836-A1.

PD 19-JUN-1997.

XX 13-DEC-1996; 96WO-US19680.

PR 14-DEC-1995; 95US-0572645.

PA (GEHO) GEN HOSPITAL CORP.

PI Fishman JA;

DR WPI: 1997-332804/30.

DR N-PSDB; AAT74884.

XX New nucleic acid from porcine retroviruses - used for detecting

PT viruses in transplant or other tissue and for assessing risk of

XX transmitting infection to graft recipient.

PS Claim 22; Fig 3; 128pp; English.

CC This is a porcine retrovirus from miniature swine containing the coding

CC region for a putative viral GAG protein. This sequence and PCR fragments

CC generated from the sequence (see AAT74812-T74882) could be used to

CC screen organs for porcine retroviruses prior to xenotransplantation.
 CC Transplantation can increase the likelihood of retroviral activation if
 CC intact and infectious proviruses are present. The porcine retroviral
 CC sequence can be used to generate probes to determine the level (e.g.
 CC copy number) of intact (i.e. potentially replicating) porcine provirus
 CC sequences in a strain of xenograft transplantation donors. It can be
 CC used to detect mutations, genetic lesions or viral recombinants and
 CC also to determine the histological localisation of activated retrovirus.
 CC using Polymerase Chain Reaction DNA Quantitation (PDQ) on blood
 CC mononuclear cells, infectivity titration and susceptibility testing can
 CC be performed. Ultimately animal donors without intact porcine retroviral
 CC sequences or a lower copy number of viral elements could be selected.

SQ Sequence 524 AA;

Query Match 29.2%; Score 74; DB 18; Length 524;
 Best Local Similarity 41.2%; Pred. No. 0.78;
 Matches 14; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

OY 12 LRDLITEAMEIKAKRHDEQORELEEEENSDDE 45
 DB 406 lrdlykaekykykreteereqrkerereee 439

RESULT 3

AAW73285 AAW73285 standard; Protein; 524 AA.

AC AAW73285;

DT 23-MAY-2001 (first entry)

DE Retroviral protein #1 found in miniature swine.

XX Retrovirus; graft transplantation; xenotransplantation; miniature swine.

XX Unidentified.

OS US6190861-B1.

PN 20-FEB-2001.

PD 13-DEC-1996; 96US-0766528.

XX 14-DEC-1995; 95US-0572645.

PA (GEHO) GEN HOSPITAL CORP.

PI Fishman JA;

DR WPI: 2001-256211/26.

DR N-PSDB; AAF77727.

XX Assessing risk of endogenous retroviruses in clinical practice and in

PT xenotransplantation, comprises using probe sequences derived from swine

XX or miniature swine retroviral genome -

PS Disclosure; Fig 3; 127pp; English.

XX The present invention relates to a method for screening a cell or tissue

CC for the presence or expression of a retrovirus (RV), comprising

CC contacting a target nucleic acid from the cell or tissue with a second

CC nucleic acid from the present invention (e.g. AAF77727 or a fragment

CC thereof). The method is useful for RV detection and to assess graft

CC transplantation risk. Screening of animals allows the elimination of

CC donors with active replication of known viruses. Inactive proviruses can

CC be detected and inactivated, allowing identification and elimination of

CC potential human pathogens derived from swine in a manner not possible in

CC the outbred human organ donor population and is important to the

CC development of human xenotransplantation.

XX Sequence 524 AA;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:14:08 ; Search time 53.55 Seconds

(without alignments)
136.575 Million cell updates/sec

Title: US-09-822-110-5

Perfect score: 253
Sequence: 1 PFKNKAPVSIILRLITRAM.....ORELEEEENDEDELDSHT 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhch:*
8: sp_mammal:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	244	96.4	321	11 09C82	09C82 mus musculu
2	244	96.4	445	11 060877	060877 mus musculu
3	244	96.4	497	11 09J110	09J110 mus musculu
4	238	94.1	491	11 054748	054748 rattus norv
5	167.5	66.2	487	11 09J111	09J111 mus musculu
6	165.5	65.4	210	4 09BR32	09BR32 homo sapien
7	90	35.6	478	5 062571	062571 suberites d
8	76.5	30.2	323	5 09V117	09V117 drosophila
9	76	30.0	642	5 09VEN3	09VEN3 drosophila
10	75	29.6	950	13 09YHC9	09YHC9 xenopus lae
11	74	29.2	524	6 062704	062704 sus scrofa
12	74	29.2	524	6 062706	062706 sus scrofa
13	73	28.9	309	5 097318	097318 plasmodium
14	73	28.9	775	11 09D6C5	09D6C5 mus musculu
15	73	28.9	775	11 09D680	09D680 mus musculu
16	72	28.5	1087	5 096923	096923 dictyosteli
17	72	28.5	1110	13 091255	091255 petromyzon
18	71.5	28.3	312	3 006859	006859 saccharomyc
19	71.5	28.3	312	3 006860	006860 saccharomyc

20	71.5	28.3	461	5 061125	061125 dictyosteli
21	71	28.1	292	10 09LVH1	09LVH1 arabidopsis
22	71	28.1	1309	5 017582	017582 caenorhabdi
23	70.5	27.9	3781	5 09VIS5	09VIS5 drosophila
24	70	27.7	775	5 09V261	09V261 drosophila
25	70	27.7	916	12 090M70	090M70 avian adeno
26	69.5	27.5	1100	5 09N330	09N330 caenorhabdi
27	69	27.3	524	6 062708	062708 sus scrofa
28	69	27.3	664	6 09N117	09N117 sus scrofa
29	69	27.3	1707	11 099K50	099K50 mus musculu
30	69	27.3	1520	4 015087	015087 homo sapien
31	69	27.3	1781	4 09UKX0	09UKX0 homo sapien
32	69	27.3	1890	4 09UKW3	09UKW3 homo sapien
33	69	27.3	2073	4 09UKM2	09UKM2 homo sapien
34	69	27.3	2376	12 09Q1X5	09Q1X5 porcine end
35	69	27.3	2376	12 09Q1X3	09Q1X3 porcine end
36	69	27.3	2378	12 09Q1X4	09Q1X4 porcine end
37	68.5	27.1	313	10 09C6H3	09C6H3 arabidopsis
38	68.5	27.1	539	5 09VPM5	09VPM5 drosophila
39	67.5	26.7	376	4 09BR60	09BR60 homo sapien
40	67.5	26.7	612	4 09BW99	09BW99 homo sapien
41	67.5	26.7	1171	3 09P3E2	09P3E2 neurospora
42	67	26.5	178	4 09UNY6	09UNY6 homo sapien
43	67	26.5	299	10 048722	048722 arabidopsis
44	67	26.5	403	5 076961	076961 necator ame
45	67	26.5	403	5 09XYF8	09XYF8 trypanosoma

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	321 AA.
ID	09C82			
AC	09C82			
DT	01-JUN-2001 (TREMBL)	17, Created		
DT	01-JUN-2001 (TREMBL)	17, Last sequence update		
DT	01-JUN-2001 (TREMBL)	17, Last annotation update		
DE	0610042106RIK	PROTEIN (FRAGMENT).		
GN	0610042106RIK			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid:10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Kasukawa T., Saito R.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,			
RA	Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Guisticich S., Hillid, Hofmann M., Hume D.A., Kamuya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	EMBL: AK002914; BAB22453.1;			
DR	MED: MGI:1914102; 0610042106RIK.			
DR	InterPro: IPR000719; Euk_Pkinase.			
DR	InterPro: IPR002290; Ser_thr_kin_actsite.			
DR	Pfam: PF00069; pkinase; 1.			

DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 FT NON_TER 1 1
 SO SEQUENCE 321 AA; 37201 MW; 2E96FD4E919F53A4 CRC64;

Query Match
 Best Local Similarity 96.4%; Score 244; DB 11; Length 321;
 Matches 48; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 50
 |||||
 DB 100 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 149

RESULT 2
 ID 060877 PRELIMINARY; PRT: 445 AA.
 060877

01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PROTEIN KINASE MESS1.
 GN STK3 OR MESS1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Han J.;
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U28726; AA075300.1;
 DR HSSP; P00518; IPRK.
 DR MGD; MGI:1928487; SEK3.
 DR InterPro; IPR000719; Euk_kinase.
 DR InterPro; IPR002290; Ser_thr_kin_actsite.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SO SEQUENCE 445 AA; 50407 MW; 7B936B28DA616F05 CRC64;

Query Match
 Best Local Similarity 96.4%; Score 244; DB 11; Length 445;
 Matches 48; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 50
 |||||
 DB 276 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 325

RESULT 3
 ID 090110 PRELIMINARY; PRT: 497 AA.
 090110

01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE STE20-LIKE KINASE MST2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Lee K.-K., Ohyama T., Yajima N., Yonehara S.;
 SO "MST", a physiological caspase substrate, highly sensitizes apoptosis

RT both upstream and downstream of caspase activation."
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF2711361; AA075790.1;
 DR InterPro; IPR000719; Euk_kinase.
 DR InterPro; IPR002290; Ser_thr_kin_actsite.
 DR InterPro; IPR001245; Tyr_Kin.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Kinase; Transferase.
 SO SEQUENCE 497 AA; 56855 MW; 9CDD365437581665 CRC64;

Query Match
 Best Local Similarity 96.4%; Score 244; DB 11; Length 497;
 Matches 48; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 50
 |||||
 DB 276 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 325

RESULT 4
 ID 054748 PRELIMINARY; PRT: 491 AA.
 054748

01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE MST2 KINASE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Aurisicchio L., Dilauro R., Zannini M.;
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; A1001529; CAA04814.1;
 DR HSSP; P00518; IPRK.
 DR InterPro; IPR000719; Euk_kinase.
 DR InterPro; IPR002290; Ser_thr_kin_actsite.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SO SEQUENCE 491 AA; 56121 MW; 90FA6B020E7FFD62 CRC64;

Query Match
 Best Local Similarity 94.1%; Score 238; DB 11; Length 491;
 Matches 46; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 50
 |||||
 DB 276 PF1KNAKPVSLRDLITFAMEIKAKRHDEQORELEEEENSDELDLSHT 325

RESULT 5
 ID 090111 PRELIMINARY; PRT: 487 AA.
 090111

01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE STE20-LIKE KINASE MST1.
 GN STK4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Db 585 EAAEAKAEVEEEAEAEVEEAEVEAEFT 617

RESULT 3

556289 regulatory protein PHO4 - yeast (Saccharomyces cerevisiae).

C:Species: Saccharomyces cerevisiae
C:Date: 02-Sep-1995 #sequence_revision 12-Apr-1996 #text_change 17-Mar-1999
C:Accession: S56289; A23482; S62245; S63839
R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sasana
Submitted to the EMBL Data Library, May 1995

A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce
A:Reference number: S56186
A:Accession: S56289
A:Molecule type: DNA
A:Residues: 1-312 <MUR>

A:Cross-references: EMBL:D50617; NID:9836685; PID:d1009914; PID:9836789; MIPS:YFR034C
R:Legrain, M.; De Wilde, M.; Hilger, F.
Nucleic Acids Res. 14, 3059-3073, 1986

A:Title: Isolation, physical characterization and expression analysis of the Saccharomy
A:Reference number: A23482; MUID:86176785
A:Accession: A23482
A:Molecule type: DNA

A:Residues: 1-289; 'RP', 292, 'RPPWRPACTSVT', 302, 'S', 304, 'T' <LEG>
A:Cross-references: EMBL:X037719
R:Murakami, Y.
Submitted to the EMBL Data Library, December 1994

A:Reference number: S62230
A:Accession: S62245
A:Molecule type: DNA
A:Residues: 1-312 <MUR>

A:Cross-references: EMBL:D44602; NID:9893419; PID:d1008636; PID:9893429
R:Exl, T.; Naitou, M.; Hagiwara, H.; Abe, M.; Ozawa, M.; Sasamura, S.I.; Sasamura, M.; T
Yeast 12, 177-190, 1996

A:Title: Fifteen open reading frames in a 30.8 kb region of the right arm of chromosome
A:Reference number: S63830; MUID:96287654
A:Accession: S63839
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-312 <EXL>
A:Cross-references: EMBL:D44602; NID:9893419; PID:d1008636; PID:9893429
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995
C:Genetics:
A:Gene: SGD:PHO4
A:Cross-references: SGD:S0001930; MIPS:YFR034C
A:Map position: 6R

Query Match 28.3%; Score 71.5; DB 2; Length 312;
Best Local Similarity 30.0%; Pred. No. 5.8;
Matches 18; Conservative 11; Mismatches 18; Indels 13; Gaps 2;

OY 2 FIKNKPVSILRDILTEAMEIKAKRHD-----EQORELEEE-----NSDEDELDS 48
Db 12 FVDDEPSSILDKGVDEFTVTKRKHGREDNFENQNDLNSQENNNSENENENENEDDS 71

RESULT 4

119170 hypothetical protein C09H6.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19170
R:Harris, B.
Submitted to the EMBL Data Library, November 1996

A:Accession: T19170
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1309 <WHL>
A:Cross-references: EMBL:Z81466; PIDN:CAB03868.1; GSPDB:GN00019; CESP:C09H6.1
A:Experimental source: clone C09H6

C:Genetics:
A:Gene: CESP:C09H6.1
A:Map position: 1
A:Introns: 107/3; 434/3; 478/3; 564/2; 1020/1; 1240/3

Query Match 28.1%; Score 71; DB 2; Length 1309;
Best Local Similarity 34.5%; Pred. No. 28;
Matches 20; Conservative 9; Mismatches 21; Indels 8; Gaps 2;

OY 1 PFIK----NAKPVSLRDILTEAM-----EIKAKRHDQORELEEEENSDEDELDSHT 50
Db 855 PEKRVQDQKRPVSLTDLNSENKMRKSTKRKMLDKVEMKGVGDEDEDSVDKGT 912

RESULT 5

150254 HMG-1 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 23-Jul-1999
C:Accession: I50254
R:Funahashi, J.; Sekido, R.; Murai, K.; Kamachi, Y.; Kondoh, H.
Development 119, 433-446, 1993

A:Title: Delta-crystallin enhancer binding protein
A:Reference number: I50222; MUID:94116444
A:Accession: I50254
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-201 <FUN>

A:Cross-references: GB:D14314; NID:9391635; PIDN:BA03260.1; PID:9391636
F:6-82/Domain: HMG box homology <HMG1>
F:89-163/Domain: HMG box homology <HMG2>

Query Match 27.3%; Score 69; DB 2; Length 201;
Best Local Similarity 44.8%; Pred. No. 6.5;
Matches 13; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 17 TEAMEIKAKRHDQORELEEEENSDEDE 45
Db 173 TKAKRKVEEEDDEDEDEDEDDDE 201

RESULT 6

522359 nonhistone chromosomal protein HMG-2a - chicken
C:Species: Gallus gallus (chicken)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 23-Jul-1999
C:Accession: S22359; S36684
R:Okazaki, T.; Endo, Y.; Ito, M.; Miyamoto, K.I.; Sasakawa, T.; Suzuki, I.; Natori, Y.
Biochim. Biophys. Acta 1130, 224-226, 1992

A:Title: Molecular cloning of chick liver HMG 2a cDNA and developmental expression of
A:Reference number: S22359; MUID:92223100
A:Accession: S22359
A:Molecule type: mRNA
A:Residues: 1-202 <OKA1>
A:Cross-references: EMBL:X63463
R:Okazaki, T.; Endo, Y.; Ito, M.; Miyamoto, K.I.; Sasakawa, T.; Suzuki, I.; Natori, Y.
Submitted to the EMBL Data Library, December 1991

A:Reference number: S36684
A:Accession: S36684
A:Molecule type: mRNA
A:Residues: 1-98; 'A', 100-202 <OKA2>

A:Cross-references: EMBL:X63463; NID:963493; PIDN:CAA5065.1; PID:963494
C:Superfamily: nonhistone chromosomal protein HMG-2; HMG box homology
C:Keywords: chromosomal protein; DNA binding; nucleus
F:6-83/Domain: HMG box homology <HMG1>
F:90-164/Domain: HMG box homology <HMG2>

Query Match

Best Local Similarity 27.3%; Score 69; DB 2; Length 202;
44.8%; Pred. No. 6.5;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:14:32; Search time 19.84 Seconds

(without alignments)
29.568 Million cell updates/sec

Title: US-09-822-110-3

Sequence: 1 EIKAKRHDEQRELEE 16

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 100059 seqs, 36664827 residues

al number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	491	1	STK3_HUMAN
2	46	56.8	487	1	STK4_HUMAN
3	45	55.6	329	1	MLO2_SCHPO
4	45	55.6	1427	1	REST_HUMAN
5	44	54.3	1537	1	MTDM_CHICK
6	42	51.9	197	1	DEMI_PHYSA
7	42	51.9	198	1	DEMI_PHYSA
8	42	51.9	198	1	DEMI_PHYSA
9	42	51.9	890	1	UZR2_HUMAN
10	42	51.9	890	1	IF2_ECOLI
11	42	51.9	897	1	IF2_ECOLI
12	42	51.9	1240	1	YF01_YEAST
13	41.5	51.2	458	1	YF01_YEAST
14	41.5	51.2	536	1	DNK1_THEMA
15	41	50.6	143	1	Y082_RICPR
16	41	50.6	180	1	GLUC_BOVIN
17	41	50.6	180	1	GLUC_BOVIN
18	41	50.6	180	1	GLUC_HUMAN
19	41	50.6	180	1	GLUC_HUMAN
20	41	50.6	180	1	GLUC_HUMAN
21	41	50.6	180	1	GLUC_HUMAN
22	41	50.6	180	1	GLUC_HUMAN
23	41	50.6	180	1	GLUC_HUMAN
24	41	50.6	180	1	GLUC_HUMAN
25	41	50.6	180	1	GLUC_HUMAN
26	41	50.6	180	1	GLUC_HUMAN
27	41	50.6	180	1	GLUC_HUMAN
28	41	50.6	180	1	GLUC_HUMAN
29	41	50.6	180	1	GLUC_HUMAN
30	41	50.6	180	1	GLUC_HUMAN
31	41	50.6	180	1	GLUC_HUMAN
32	41	50.6	180	1	GLUC_HUMAN
33	41	50.6	180	1	GLUC_HUMAN

34	39	48.1	444	1	STU1_YEAST
35	39	48.1	849	1	PRO5_YEAST
36	39	48.1	892	1	IF2_SALTY
37	39	48.1	945	1	PNP2_YEAST
38	39	48.1	1010	1	WNT5_DROME
39	39	48.1	1052	1	BU1B_MOUSE
40	39	48.1	1394	1	E75B_DROME
41	39	48.1	2230	1	G0G4_HUMAN
42	38.5	47.5	427	1	YF09_MYCPN
43	38.5	47.5	539	1	PNUT_DROME
44	38.5	47.5	614	1	RA26_SCHPO
45	38	46.9	336	1	RPOA_THEMA

ALIGNMENTS

RESULT	ID	STANDARD	PRT	491 AA
1	STK3_HUMAN	013188; Q15801; Q15445;		
AC	013188; Q15801; Q15445;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SERINE/THREONINE PROTEIN KINASE 3 (EC 2.7.1.37) (STE20-LIKE KINASE			
DE	MST2) (MST-2) (MAMMALIAN STE20-LIKE PROTEIN KINASE 2)			
DE	(SERINE/THREONINE PROTEIN KINASE KRS-1).			
GN	STK3 OR MST2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI:Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-96144292; PubMed-8566796;			
RA	Creasy C.L., Chernoff J.,			
RT	"Cloning and characterization of a member of the MST subfamily of			
RT	Ste20-like kinases."			
RL	Gene 167:303-306(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-96413604; PubMed-8816758;			
RA	Taylor L.K., Wang H.C., Erikson R.L.,			
RT	"Newly identified stress-responsive protein kinases, KRS-1 and KRS-2."			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996)			
RN	[3]			
RP	SEQUENCE OF 96-203 FROM N.A.			
RX	MEDLINE-94100173; PubMed-8274451;			
RA	Schultz S.J., Nigg E.A.,			
RT	"Identification of 21 novel human protein kinases, including 3 members			
RT	of a family related to the cell cycle regulator nimA of Aspergillus			
RT	nidulans."			
RL	Cell Growth Differ. 4:821-830(1993).			
CC	FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT			
CC	MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY			
CC	SIMILARITY).			
CC	CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.			
CC	SUBCELLULAR LOCATION: CYTOSOL; PLASMA MEMBRANE (BY SIMILARITY).			
CC	TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,			
CC	SKELETAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT			
CC	HEART, LUNG AND BRAIN TISSUES.			
CC	SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	STE20 SUBFAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

DR EMBL: U26424; AAC50386.1; -
 DR EMBL: U60206; AAB17261.1; -
 DR EMBL: Z25422; CAAB0909.1; -
 DR HSSP: P00518; 1PHK.
 DR MIM: 605030; -
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR Pfam: PF00069; Pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST_FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
 DR Trasnferase: Serine/threonine-protein kinase; ATP-binding.
 DR DOMAIN: 27 278 PROTEIN KINASE.
 DR NP_BIND: 33 41 ATP (BY SIMILARITY).
 DR BINDING: 56 56 ATP (BY SIMILARITY).
 DR ACT_SITE: 146 146 BY SIMILARITY.
 DR DOMAIN: 308 314 POLY-GLU.
 DR CONFLICT: 96 98 WIV -> YLY (IN REF. 3).
 DR CONFLICT: 121 121 D -> Y (IN REF. 3).
 DR CONFLICT: 203 203 D -> G (IN REF. 3).
 DR CONFLICT: 303 303 D -> E (IN REF. 2).
 DR CONFLICT: 332 334 GEC -> ESV (IN REF. 2).
 DR CONFLICT: 332 334
 DR SEQUENCE 491 AA; 56261 MW; 9CA3B0644FC14A9 CRC64;

Query Match 100.0%; Score 81; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 7; 6e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIKAKRHDEOORELEE 16
 |||||||
 Db 296 EIKAKRHDEOORELEE 311

RESULT 2
 ID STK4_HUMAN STANDARD; PRT; 487 AA.
 AC Q13043; Q15802; Q9NTZ4;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SERINE/THREONINE PROTEIN KINASE 4 (EC 2.7.1.37) (STE20-LIKE KINASE
 DE MST1) (MST-1) (MAMMALIAN STE20-LIKE PROTEIN KINASE 1)
 DE (SERINE/THREONINE PROTEIN KINASE KRS-2).
 DE STK4 OR MST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394929; Pubmed=7665586;
 RA Creasy C.L., Chernoff J.;
 RT "Cloning and characterization of a human protein kinase with homology
 RT to Ste20."
 RL J. Biol. Chem. 270:21695-21700(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413604; Pubmed=8816758;
 RA Taylor L.R., Wang H.C., Erikson R.L.;
 RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
 RT 2."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
 RN [3]
 RP SEQUENCE OF 1-435 FROM N.A.
 RA Laird G.;
 CC CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
 CC ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
 CC THE KINASE ACTIVITY.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: UNOBTAINABLE EXPRESSED.
 CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC STE20 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U18297; AAB83254.1; -
 DR EMBL: U60207; AAB17262.1; -
 DR EMBL: AL109839; CAB89421.1; -
 DR HSSP: P24941; IMCL.
 DR MIM: 604965; -
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR Pfam: PF00069; Pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST_FALSE_NEG.
 DR Trasnferase: Serine/threonine-protein kinase; ATP-binding.
 DR DOMAIN: 30 281 PROTEIN KINASE.
 DR NP_BIND: 36 44 ATP (BY SIMILARITY).
 DR BINDING: 59 59 ATP (BY SIMILARITY).
 DR ACT_SITE: 149 149 BY SIMILARITY.
 DR DOMAIN: 373 378 POLY-GLU.
 DR CONFLICT: 222 222 P -> R (IN REF. 1).
 DR CONFLICT: 312 312 V -> M (IN REF. 1).
 DR SEQUENCE 487 AA; 55630 MW; 150758EBC5F77D5C CRC64;

Query Match 56.88; Score 46; DB 1; Length 487;
 Best Local Similarity 43.88; Pred. No. 10;
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 EIKAKRHDEOORELEE 16
 |||||
 Db 299 DVKLKROESQOREVDQ 314

RESULT 3
 ID MLO2_SCHPO STANDARD; PRT; 329 AA.
 AC Q09329;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROTEIN MLO2.
 DE MLO2 OR SPBC4.05.
 GN Schizosaccharomyces pombe (fission yeast).
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97128260; Pubmed=8972853;
 RA Javerzat J.-P., Cranston G., Allshire R.A.;
 RT "Fission yeast genes which disrupt mitotic chromosome segregation
 RT when overexpressed."
 RL Nucleic Acids Res. 24:4676-4683(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:14:07 ; Search time 53.55 Seconds
(without alignments)
43.704 Million cell updates/sec

Title: US-09-822-110-3
Perfect score: 81
Sequence: 1 EIKAKRHDEQRELEE 16

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_17.*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	77	95.1	321	11	Q9CW82 mus musculus
2	77	95.1	445	11	Q60877 mus musculus
3	77	95.1	497	11	Q9J110 mus musculus
4	74	91.4	491	11	Q54748 mus musculus
5	48	59.3	799	4	Q60302 mus musculus
6	47	58.0	2357	5	Q901M8 mus musculus
7	46.5	57.4	222	2	Q9R1D9 mus musculus
8	46.5	57.4	233	2	Q68746 mus musculus
9	46.5	57.4	210	4	Q9BR32 mus musculus
10	45	55.6	487	11	Q9J111 mus musculus
11	45	55.6	565	5	Q9VWD4 mus musculus
12	45	55.6	625	10	Q9SP13 mus musculus
13	45	55.6	666	10	Q9SP15 mus musculus
14	45	55.6	666	10	Q9SP14 mus musculus
15	45	55.6	1320	11	Q9J125 mus musculus
16	44.5	54.9	3328	5	Q9VUB5 mus musculus
17	44.5	54.9	455	3	Q12080 mus musculus
18	44	54.3	437	11	Q9CU41 mus musculus
19	44	54.3	1621	5	Q01900 mus musculus

20	43	53.1	344	5	Q76741 dictyostell
21	43	53.1	363	4	Q9NV16 dictyostell
22	43	53.1	444	2	Q9RV42 dictyostell
23	43	53.1	477	5	Q9N4J9 dictyostell
24	43	53.1	499	10	Q23211 dictyostell
25	43	53.1	522	10	Q9FX85 dictyostell
26	43	53.1	643	10	Q39846 dictyostell
27	43	53.1	1011	11	Q9NMR1 dictyostell
28	43	53.1	1131	5	Q9VGE4 dictyostell
29	43	53.1	1219	5	Q9VYJ9 dictyostell
30	43	53.1	1645	5	Q9U263 dictyostell
31	43	53.1	1655	5	Q44498 dictyostell
32	43	53.1	2785	11	Q9NMR3 dictyostell
33	43	53.1	3259	4	Q14789 dictyostell
34	42.5	52.5	670	4	Q9NMQ3 dictyostell
35	42.5	52.5	682	4	Q9NX83 dictyostell
36	42.5	52.5	746	4	Q9HA51 dictyostell
37	42.5	52.5	975	4	Q9H9Y7 dictyostell
38	42	51.9	251	11	Q9CSR1 dictyostell
39	42	51.9	254	6	Q9GM10 dictyostell
40	42	51.9	419	3	Q74786 dictyostell
41	42	51.9	438	6	Q9BH01 dictyostell
42	42	51.9	488	5	Q9GPA2 dictyostell
43	42	51.9	593	10	Q9MAB4 dictyostell
44	42	51.9	726	2	Q9EUZ8 dictyostell
45	42	51.9	727	2	Q9EUZ4 dictyostell

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	321-AA
Q9CW82	Q9CW82	01-JUN-2001 (TRENDEL, 17, Created)		
Q9CW82	Q9CW82	01-JUN-2001 (TRENDEL, 17, Last sequence update)		
Q9CW82	Q9CW82	01-JUN-2001 (TRENDEL, 17, Last annotation update)		
Q9CW82	Q9CW82	0610042106R1K PROTEIN (FRAGMENT)		
Q9CW82	Q9CW82	0610042106R1K		
Q9CW82	Q9CW82	Mus musculus (Mouse)		
Q9CW82	Q9CW82	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
Q9CW82	Q9CW82	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
Q9CW82	Q9CW82	NCBI:TaxID=10090;		
Q9CW82	Q9CW82	[1]		
Q9CW82	Q9CW82	SEQUENCE FROM N.A.		
Q9CW82	Q9CW82	STRAIN=C57BL/6J; TISSUE=KIDNEY;		
Q9CW82	Q9CW82	MEDLINE=21085660; PubMed=11217851;		
Q9CW82	Q9CW82	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
Q9CW82	Q9CW82	Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,		
Q9CW82	Q9CW82	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,		
Q9CW82	Q9CW82	Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,		
Q9CW82	Q9CW82	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
Q9CW82	Q9CW82	Fleischmann W., Gaasterland T., Gissi C., King B., Kocula H.,		
Q9CW82	Q9CW82	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,		
Q9CW82	Q9CW82	Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,		
Q9CW82	Q9CW82	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
Q9CW82	Q9CW82	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
Q9CW82	Q9CW82	Brownstein M.J., Buft C., Fletcher C., Fujita M., Gariboldi M.,		
Q9CW82	Q9CW82	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
Q9CW82	Q9CW82	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbarts P.,		
Q9CW82	Q9CW82	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
Q9CW82	Q9CW82	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
Q9CW82	Q9CW82	Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,		
Q9CW82	Q9CW82	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,		
Q9CW82	Q9CW82	Hayashizaki Y.,		
Q9CW82	Q9CW82	"Functional annotation of a full-length mouse cDNA collection."		
Q9CW82	Q9CW82	Nature 409:685-690(2001).		
Q9CW82	Q9CW82	EMBL: AK002914; BAB22453.1;		
Q9CW82	Q9CW82	MDG: MGI:1914102; 0610042106R1K.		
Q9CW82	Q9CW82	InterPro: IPR000719; Euk_kinase.		
Q9CW82	Q9CW82	InterPro: IPR002290; Ser-thr_kin_actsite.		
Q9CW82	Q9CW82	Pfam: PF00069; pkinase; I.		

DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 FT NON_TER
 SQ SEQUENCE 321 AA; 37201 MW; 2E96FDAE919F53A4 CRC64;

Query Match 95.1%; Score 77; DB 11; Length 321;
 Best Local Similarity 93.8%; Pred. No. 0.0018;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIKAKRHDEQRELEF 16
 DB 120 EIKAKRHDEQRELEF 135

RESULT 2
 060877 PRELIMINARY; PRT; 445 AA.

DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DE PROTEIN_KINASE_MESS1.
 GN STK3 OR MESS1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Han J.
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U28726; AAA75300.1; -
 DR HSP; P00518; IPIK.
 DR MGD; MGI:1928487; STK3.
 DR InterPro: IPR000719; Ser thr_kin_actsite.
 DR InterPro: IPR002290; Ser thr_kin_actsite.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 445 AA; 50407 MW; 7B936B28DA616F05 CRC64;

Query Match 95.1%; Score 77; DB 11; Length 445;
 Best Local Similarity 93.8%; Pred. No. 0.0024;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIKAKRHDEQRELEF 16
 DB 296 EIKAKRHDEQRELEF 311

RESULT 3
 09J110 PRELIMINARY; PRT; 497 AA.

AC 09J110;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE STE20-LIKE KINASE MST2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RA Lee K.-K., Ohyama T., Yajima N., Yonehara S.;
 FT "MST, a physiological caspase substrate, highly sensitizes apoptosis

RT both upstream and downstream of caspase activation";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF271361; AAF75790.1; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR InterPro: IPR001245; Tyr_kin.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 497 AA; 56855 MW; 9CDD365437581665 CRC64;

Query Match 95.1%; Score 77; DB 11; Length 497;
 Best Local Similarity 93.8%; Pred. No. 0.0026;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIKAKRHDEQRELEF 16
 DB 296 EIKAKRHDEQRELEF 311

RESULT 4
 054748 PRELIMINARY; PRT; 491 AA.

ID 054748;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DE MST2 KINASE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Aurisicchio L., Dilauro R., Zannini M.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ001529; CAA04814.1; -
 DR HSP; P00518; IPIK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser thr_kin_actsite.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 491 AA; 56121 MW; 90FA6B020E7FFD62 CRC64;

Query Match 91.4%; Score 74; DB 11; Length 491;
 Best Local Similarity 87.5%; Pred. No. 0.0069;
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIKAKRHDEQRELEF 16
 DB 296 EIKAKRHDEQRELEF 311

RESULT 5
 060302 PRELIMINARY; PRT; 799 AA.

AC 060302;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE KIA0555 PROTEIN.
 GN KIA0555.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:04:58 ; Search time 26.93 Seconds
(without alignments)
13.370 Million cell updates/sec

Title: US-09-822-110-3

Perfect score: 81

Sequence: 1 EIKAKRHDEQRELEE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfills1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	58.0	487	2	US-08-712-709-8
2	47	58.0	487	3	US-09-111-444-8
3	47	58.0	487	4	US-09-541-228-8
4	41	50.6	180	3	US-08-784-582-56
5	41	50.6	180	3	US-08-784-582-58
6	41	50.6	180	3	US-08-784-582-61
7	41	50.6	360	3	US-08-784-582-73
8	41	50.6	737	1	US-08-188-582-16
9	41	50.6	737	1	US-08-646-715-16
10	40	49.4	643	2	US-08-551-356-4
11	40	49.4	643	5	PCT-US93-12687-4
12	40	49.4	644	1	US-08-206-176-2
13	38.5	47.5	539	2	US-08-978-182-3
14	38.5	47.5	539	2	US-09-205-681-3
15	38	46.9	186	4	US-09-446-504-70
16	38	46.9	450	2	US-08-665-037-2
17	38	46.9	450	2	US-08-666-067-2
18	38	46.9	450	2	US-08-732-870-2
19	38	46.9	926	3	US-08-755-587-187
20	38	46.9	1964	3	US-08-836-325-10
21	38	46.9	1989	3	US-08-836-325-12
22	38	46.9	2548	4	US-09-172-422-1
23	37	45.7	35	4	US-09-248-588-26
24	37	45.7	203	1	US-08-216-593-5
25	37	45.7	203	5	PCT-US93-12380-5
26	37	45.7	205	1	US-07-992-827D-5
27	37	45.7	209	1	US-07-992-827D-2

28	37	45.7	209	1	US-08-216-593-2	Sequence 2, App11
29	37	45.7	209	5	PCT-US93-12380-2	Sequence 2, App11
30	37	45.7	288	3	US-08-312-949-4	Sequence 4, App11
31	37	45.7	288	3	US-08-446-201-4	Sequence 4, App11
32	37	45.7	289	1	US-08-072-070-4	Sequence 4, App11
33	37	45.7	289	1	US-08-469-434-4	Sequence 4, App11
34	37	45.7	289	1	US-08-214-222-4	Sequence 4, App11
35	37	45.7	289	2	US-08-467-852A-5	Sequence 5, App11
36	37	45.7	289	2	US-08-247-491A-5	Sequence 5, App11
37	37	45.7	323	1	US-07-992-827D-1	Sequence 1, App11
38	37	45.7	323	1	US-08-216-593-1	Sequence 1, App11
39	37	45.7	323	5	PCT-US93-12380-1	Sequence 1, App11
40	37	45.7	329	4	US-08-884-681-3	Sequence 3, App11
41	37	45.7	329	4	US-09-258-643-3	Sequence 3, App11
42	37	45.7	372	4	US-08-884-681-1	Sequence 1, App11
43	37	45.7	372	4	US-09-258-643-1	Sequence 1, App11
44	37	45.7	619	1	US-08-465-746-2	Sequence 2, App11
45	37	45.7				

ALIGNMENTS

RESULT 1
US-08-712-709-8
Sequence 8, Application US/08712709
Patent No. 5863780
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,709
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791
US-08-712-709-8

Query Match 58.0%; Score 47; DB 2; Length 487;
Best Local Similarity 43.8%; Pred. No. 7.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 1 EIKAKRHDEQRELEE 16

Db 299 DVKLKROESQOREMDQ 314

RESULT 2

US-09-111-444-8

Sequence 8, Application US/09111444

Patent No. 6045792

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/111,444

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/712,709

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 487 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1117791

US-09-111-444-8

Query Match 58.0%; Score 47; DB 3; Length 487;

Best Local Similarity 43.8%; Pred. No. 7.3;

Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 EIKAKRHDEQORELEE 16

Db 299 DVKLKROESQOREMDQ 314

RESULT 3

US-09-541-228-8

Sequence 8, Application US/09541228

Patent No. 6232077

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/541,228

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/712,709

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 487 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1117791

US-09-541-228-8

Query Match 58.0%; Score 47; DB 4; Length 487;

Best Local Similarity 43.8%; Pred. No. 7.3;

Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 1 EIKAKRHDEQORELEE 16

Db 299 DVKLKROESQOREMDQ 314

RESULT 4

US-08-784-582-56

Sequence 56, Application US/08784582

Patent No. 6110707

GENERAL INFORMATION:

APPLICANT: Newgard, Christopher B.

APPLICANT: Halban, Philippe A.

APPLICANT: No. 6110707Mington, Karl D.

APPLICANT: Clark, Samuel A.

APPLICANT: Thijsen, Anice E.

APPLICANT: Quade, Christian

APPLICANT: Kruse, Fred

APPLICANT: McGarry, Dennis

TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM

TITLE OF INVENTION: SECRETORY CELL LINES

NUMBER OF SEQUENCES: 79

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:05:59 ; Search time 54.15 seconds

(without alignments)
21.887 Million cell updates/sec

Title: US-09-822-110-3

Sequence: 1 EIKARKHDEQORELEE 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

A_Geneseq_1101:*

1: /SIDSB8/gcgdata/geneseq/geneseq/AA1980.DAT:*

2: /SIDSB8/gcgdata/geneseq/geneseq/AA1981.DAT:*

3: /SIDSB8/gcgdata/geneseq/geneseq/AA1982.DAT:*

4: /SIDSB8/gcgdata/geneseq/geneseq/AA1983.DAT:*

5: /SIDSB8/gcgdata/geneseq/geneseq/AA1984.DAT:*

6: /SIDSB8/gcgdata/geneseq/geneseq/AA1985.DAT:*

7: /SIDSB8/gcgdata/geneseq/geneseq/AA1986.DAT:*

8: /SIDSB8/gcgdata/geneseq/geneseq/AA1987.DAT:*

9: /SIDSB8/gcgdata/geneseq/geneseq/AA1988.DAT:*

10: /SIDSB8/gcgdata/geneseq/geneseq/AA1989.DAT:*

11: /SIDSB8/gcgdata/geneseq/geneseq/AA1990.DAT:*

12: /SIDSB8/gcgdata/geneseq/geneseq/AA1991.DAT:*

13: /SIDSB8/gcgdata/geneseq/geneseq/AA1992.DAT:*

14: /SIDSB8/gcgdata/geneseq/geneseq/AA1993.DAT:*

15: /SIDSB8/gcgdata/geneseq/geneseq/AA1994.DAT:*

16: /SIDSB8/gcgdata/geneseq/geneseq/AA1995.DAT:*

17: /SIDSB8/gcgdata/geneseq/geneseq/AA1996.DAT:*

18: /SIDSB8/gcgdata/geneseq/geneseq/AA1997.DAT:*

19: /SIDSB8/gcgdata/geneseq/geneseq/AA1998.DAT:*

20: /SIDSB8/gcgdata/geneseq/geneseq/AA1999.DAT:*

21: /SIDSB8/gcgdata/geneseq/geneseq/AA2000.DAT:*

22: /SIDSB8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	59.3	62	22	AA18700
2	45	55.6	487	20	AA121674
3	45	55.6	623	19	AA162830
4	45	55.6	666	19	AA162828
5	45	55.6	666	19	AA162829
6	45	55.6	1392	20	AA166999
7	45	55.6	1427	12	AA110534
8	43	53.1	89	21	AA162274
9	43	53.1	100	21	AA162274
10	43	53.1	363	22	AA183034
11	42.5	52.5	243	21	AA180619

12	42.5	52.5	746	22	AA183938	Human protein sequ
13	42.5	52.5	975	22	AA184042	Human protein sequ
14	42	51.9	164	22	AA183548	Human gastric can
15	42	51.9	283	22	AA183799	Human colon cancer
16	42	51.9	482	20	AA187067	Renal cancer assoc
17	42	51.9	752	18	AA184178	Human transcriptio
18	42	51.9	898	20	AA185936	Human SUI1 protei
19	42	51.9	898	20	AA185936	Human SUI1 serine/
20	42	51.9	898	21	AA184244	Human cell signal
21	42	51.9	899	21	AA184391	Human cell signal
22	42	51.9	1100	22	AA183286	Human ORF2955
23	42	51.9	1299	21	AA185863	Human protein sequ
24	41	50.6	121	22	AA183088	Protein regulating
25	41	50.6	156	22	AA182079	S. epidermidis ope
26	41	50.6	180	18	AA182079	Angiotensin conver
27	41	50.6	180	18	AA182079	Rat preproglucagon
28	41	50.6	180	18	AA182080	Human preproglucag
29	41	50.6	180	18	AA182081	Human preproglucag
30	41	50.6	180	18	AA18384	Rat prepro-glucago
31	41	50.6	180	20	AA183812	Preproglucagon pro
32	41	50.6	180	21	AA186773	Rat preproglucagon
33	41	50.6	180	21	AA186773	Human preproglucag
34	41	50.6	180	21	AA186775	Mutant human prepr
35	41	50.6	180	21	AA186777	Human growth hormo
36	41	50.6	180	21	AA186777	Human polypeptide
37	41	50.6	180	21	AA186777	Human polypeptide
38	41	50.6	180	21	AA186777	S cerevisiae apopt
39	41	50.6	180	21	AA186777	TATA-binding prote
40	41	50.6	180	21	AA186777	Human TATA-binding
41	40.5	50.0	180	21	AA186777	TATA-binding prote
42	40.5	50.0	180	21	AA186777	Leucocytoczoan prot
43	40	49.4	180	21	AA186777	Leucocytoczoan prot
44	40	49.4	180	21	AA186777	Environmental stre
45	40	49.4	180	21	AA186777	Human pancreatic c

ALIGNMENTS

RESULT 1	AA18700	standard; Protein: 62 AA.
ID	AA18700	
AC	AA18700	
DT	12-OCT-2001	(first entry)
DE	Peptide #5134 encoded by probe for measuring cervical gene expression.	
XX	Probe: human: microarray; gene expression; cervical epithelial cell;	
KW	cervical cancer.	
OS	Homo sapiens.	
PN	W0200157278-A2.	
PD	09-AUG-2001.	
PF	30-JAN-2001; 2001WO-US00670.	
PR	04-FEB-2000; 2000US-0180312.	
PR	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
PR	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
PR	27-SEP-2000; 2000US-0236359.	
PR	04-OCT-2000; 2000GB-0024263.	
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX	WPI; 2001-488901/53.	

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells
 PS
 XX Claim 27; SEQ ID No 23526; 487bp; English.
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENP: see AAI10068-AI128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 62 AA;
 OY 1 EIKAKRHDEQORELE 15
 DB 29 elkakheekkelq 43
 RESULT 2
 AAY21674 standard; Protein; 487 AA.
 XX ID AAY21674;
 AC 18-AUG-1999 (first entry)
 DT
 XX Human Ste20 homologue polypeptide S201.
 DE
 XX Sterile 20; serine/threonine kinase; C12.2b5; 5e.new; murine;
 KW erythroleukemia; mitogenic signaling; cell differentiation; metastasis;
 KW signaling pathway; cancer; ischemic stroke; heart disease; inflammation.
 XX
 XX Homo sapiens.
 OS
 XX MO9929857-AI.
 XX 17-JUN-1999.
 XX
 XX 09-DEC-1998; 98WO-US26116.
 PF
 XX 09-DEC-1997; 97US-0069078.
 PR
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 PA
 PI Agarwal S, Best J, Vail B, Zon LI;
 DR WPI; 1999-385605/32.
 XX
 XX Murine homologues of the sterile 20 family of serine/threonine
 PT kinases, useful in elucidation of mitogenic signalling pathways
 PT
 XX Disclosure; Fig 4; 37pp; English.
 PS
 XX The invention provides two members of the sterile 20 (ste20) family of
 CC serine/threonine kinases (C12.2b5 and 5e.new). Host cells transfected
 CC with vectors comprising the nucleic acids encoding the polypeptides are
 CC used for their recombinant expression. The Ste20 kinases, isolated from a
 CC murine erythroleukemia cDNA library, are regulatory molecules involved in
 CC mitogenic signaling as well as other cellular phenomena such as
 CC morphology and motility. These activities are important factors in
 CC development, cell differentiation, cancer and metastases. Manipulation of

CC the C12.2b5 and 5e.new proteins or polynucleotides are useful for
 CC manipulation of the signaling pathways involved and will allow
 CC development of reagents to modulate these signaling pathways. The kinases
 CC are also important for ischemic stroke, heart disease and inflammation.
 CC The antibodies can be used to inhibit the activity of the kinase,
 CC especially in vitro and in cell extracts, as well as identify the kinases
 CC in immunoassays. The present sequence represents a S201 polypeptide, a
 CC human homologue of Ste20.
 CC
 XX Sequence 487 AA;
 OY 1 EIKAKRHDEQORELE 16
 DB 299 dvklkrqeqagrevdq 314
 RESULT 3
 AAW62830 standard; Protein; 625 AA.
 XX ID AAW62830;
 AC 27-OCT-1998 (first entry)
 DT
 XX Macadamia integrifolia antimicrobial protein.
 DE
 XX antimicrobial protein; infestation; control.
 KW
 XX Macadamia integrifolia.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..28
 FT /note="signal peptide"
 FT 29..666
 FT Protein /note="mature protein"
 XX
 XX WO9827805-AI.
 PN
 XX 02-JUL-1998.
 PD
 XX 22-DEC-1997; 97WO-AU00874.
 PE
 XX 20-DEC-1996; 96AU-0004275.
 PR
 XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY;
 PA Bower NI, Goulier KC, Green JL, Manners JM, Marcus JP;
 XX WPI; 1998-377279/32.
 DR N-PSDB; AAV42316.
 DR
 XX Novel anti-microbial protein from e.g. Macadamia integrifolia
 PT useful for controlling microbial infestations of plants or mammals
 PT
 XX Claim 1; Page 43-45; 96pp; English.
 PS
 XX The sequence is that of an antimicrobial protein which can
 CC be used to control microbial infestations in plants and mammalian
 CC animals.
 CC
 XX Sequence 625 AA;
 OY 5 KRHDQORELE 16
 Query Match 55.6%; Score 45; DB 19; Length 625;
 Best Local Similarity 75.0%; Pred. No. 45;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:08:32 ; Search time 12.48 Seconds
(without alignments)
28.850 Million cell updates/sec

Title: US-09-822-110-3

Sequence: 1 EIKAKRHDEQRELE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

1 number of hits satisfying chosen parameters: 96934

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	34.6	12	2	US-08-973-563A-29
2	28	34.6	12	2	US-08-973-559-29
3	28	34.6	12	2	US-08-973-235-9
4	28	34.6	12	4	US-09-051-986-12
5	28	34.6	14	1	US-07-686-322A-6
6	28	34.6	14	1	US-08-002-999-6
7	28	34.6	12	4	US-09-051-986-13
8	27	33.3	12	4	US-08-199-776-24
9	27	33.3	12	2	US-08-973-563A-31
10	27	33.3	12	2	US-08-973-559-31
11	27	33.3	12	3	US-08-663-731-24
12	27	33.3	12	3	US-08-879-338-24
13	27	33.3	12	3	US-08-993-235-11
14	27	33.3	12	5	PCT-US95-02044-24
15	27	33.3	13	4	US-09-331-507A-10
16	27	33.3	13	4	US-09-331-507A-25
17	27	33.3	14	1	US-08-199-776-18
18	27	33.3	14	1	US-08-352-179-27
19	27	33.3	14	3	US-08-663-731-18
20	27	33.3	14	3	US-08-879-338-18
21	27	33.3	14	5	PCT-US95-02044-18
22	27	33.3	15	1	US-08-199-776-20
23	27	33.3	15	3	US-08-663-731-20
24	27	33.3	15	3	US-08-879-338-20
25	27	33.3	15	5	PCT-US95-02044-20
26	27	33.3	16	1	US-08-126-564A-36
27	27	33.3	16	5	PCT-US94-09143-36

28	26	32.1	11	6	5210075-10	Patent No. 5210075
29	26	32.1	12	4	US-09-518-046-34	Sequence 34, Appl
30	26	32.1	13	4	US-07-987-272A-6	Sequence 6, Appl
31	26	32.1	14	1	US-08-199-776-19	Sequence 19, Appl
32	26	32.1	14	3	US-08-663-731-19	Sequence 19, Appl
33	26	32.1	14	3	US-08-879-338-19	Sequence 19, Appl
34	26	32.1	14	5	PCT-US95-02044-19	Sequence 19, Appl
35	26	32.1	15	1	US-08-287-717-7	Sequence 7, Appl
36	26	32.1	15	1	US-08-481-888A-17	Sequence 17, Appl
37	26	32.1	15	1	US-08-485-273A-17	Sequence 17, Appl
38	26	32.1	15	1	US-08-441-914-7	Sequence 7, Appl
39	26	32.1	15	2	US-08-973-563A-17	Sequence 17, Appl
40	26	32.1	15	2	US-08-973-559-17	Sequence 17, Appl
41	25	30.9	9	5	PCT-US95-03236-56	Sequence 56, Appl
42	25	30.9	10	1	US-08-481-888A-21	Sequence 21, Appl
43	25	30.9	10	1	US-08-485-273A-21	Sequence 21, Appl
44	25	30.9	10	2	US-08-973-563A-21	Sequence 21, Appl
45	25	30.9	10	2	US-08-973-559-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-973-563A-29
Sequence 29, Application US/08973563A
Patent No. 5885965
GENERAL INFORMATION:
APPLICANT: Oppenheim, Frank G.
APPLICANT: Xu, Tao
APPLICANT: Spacciapoli, Peter
APPLICANT: Roberts, F. D.
APPLICANT: Friden, Phillip M.
TITLE OF INVENTION: Anti-Fungal D-Amino Acid Histatin-Based Peptides
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: US
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,563A
FILING DATE: 07-JUN-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,273
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: PER95-02A2
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
INFORMATION FOR SEQ. ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 1..12
OTHER INFORMATION: /note= "At least one amino acid

OTHER INFORMATION: must have a D configuration."
US-08-973-563A-29

Query Match 34.6%; Score 28; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 AKRHDEQORE 13
||||| : : : :
DB 1 AKRHHKVKRK 10

RESULT 2
US-08-973-559-29
Sequence 29, Application US/08973559
Patent No. 5912230
GENERAL INFORMATION:

APPLICANT: OPPENHEIM, FRANK G.
APPLICANT: XU, TAO
APPLICANT: ROBERTS, F. D.
APPLICANT: SPACCIAPOLI, PETER
APPLICANT: FRIDEN, PHILIP M.
TITLE OF INVENTION: Anti-Fungal and Anti-Bacterial
TITLE OF INVENTION: Histatin-Based Peptides
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millita Drive
CITY: Lexington
STATE: MA
COUNTRY: US
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,559
FILING DATE: 07-JUN-1996
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,888
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.

REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: PER95-01A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-973-559-29

Query Match 34.6%; Score 28; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 AKRHDEQORE 13
||||| : : : :
DB 1 AKRHHKVKRK 10

RESULT 3
US-08-993-235-9

Sequence 9, Application US/08993235
Patent No. 6084064
GENERAL INFORMATION:

APPLICANT: OPPENHEIM, FRANK G.
APPLICANT: ROBERTS, F. DONALD
APPLICANT: XU, TAO
APPLICANT: SPACCIAPOLI, PETER
TITLE OF INVENTION: ANTIFUNGAL AND ANTIBACTERIAL PEPTIDES
FILE REFERENCE: 50032/002001
CURRENT APPLICATION NUMBER: US/08/993,235
CURRENT FILING DATE: 1997-12-18
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9

LENGTH: 12
TYPE: PRT
ORGANISM: Homo sapiens
US-08-993-235-9

Query Match 34.6%; Score 28; DB 3; Length 12;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 AKRHDEQORE 13
||||| : : : :
DB 1 AKRHHKVKRK 10

RESULT 4
US-09-051-986-12
Sequence 12, Application US/09051986A
Patent No. 6191113
GENERAL INFORMATION:

APPLICANT: NAKAHARA, YO
APPLICANT: HARA, Saburo
APPLICANT: KAMIKUBO, Yutchi
APPLICANT: TAKEMOTO, Sumiyo
APPLICANT: MIYAMOTO, Seiji
TITLE OF INVENTION: NOVEL PEPTIDE
FILE REFERENCE: NAKAHARA-1

CURRENT APPLICATION NUMBER: US/09/051,986A
CURRENT FILING DATE: 1998-04-24
EARLIER APPLICATION NUMBER: JP 300792/1995
EARLIER FILING DATE: 1995-10-24
EARLIER APPLICATION NUMBER: PCT/JP96/03080
EARLIER FILING DATE: 1996-10-23
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12

LENGTH: 12
TYPE: PRT
ORGANISM: Human Tissue Factor Pathway Inhibitor
US-09-051-986-12

Query Match 34.6%; Score 28; DB 4; Length 12;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 IKAKRHDEQOR 12
||| : : : :
DB 1 IKTKRRKKRKR 11

RESULT 5
US-07-686-322A-6
Sequence 6, Application US/07686322A
Patent No. 5312733
GENERAL INFORMATION:

APPLICANT: Macleod Dr., Carol L.
TITLE OF INVENTION: NO. 5312733el T-cell Lymphoma cDNA Clones
NUMBER OF SEQUENCES: 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:06:07 ; Search time 23.79 Seconds
(without alignments)
49.818 Million cell updates/sec

Title: US-09-822-110-3
Perfect score: 81
Sequence: I:EIKAKRHDEQRELEE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

1 number of hits satisfying chosen parameters: 178144

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID58/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SID58/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SID58/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SID58/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SID58/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SID58/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SID58/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SID58/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SID58/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SID58/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SID58/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SID58/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SID58/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SID58/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SID58/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SID58/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SID58/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SID58/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SID58/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SID58/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SID58/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SID58/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	40.7	9	16	AAW21326
2	33	40.7	14	22	AAW21326
3	31	38.3	11	16	AAW21333
4	29	35.8	14	19	AAW10852
5	29	35.8	15	21	AAW10852
6	28	34.6	10	19	AAW46263
7	28	34.6	12	18	AAW21922
8	28	34.6	12	18	AAW19465
9	28	34.6	12	18	AAW19419
10	28	34.6	12	20	AAW25145
11	28	34.6	12	21	AAW11979

12	28	34.6	14	15	AAW3469
13	28	34.6	14	18	AAW21923
14	27	33.3	9	22	AAW83807
15	27	33.3	10	18	AAW42713
16	27	33.3	10	22	AAW3166
17	27	33.3	10	22	AAW3168
18	27	33.3	11	22	AAW83811
19	27	33.3	11	20	AAW02798
20	27	33.3	12	16	AAW82654
21	27	33.3	12	18	AAW19467
22	27	33.3	12	18	AAW19421
23	27	33.3	12	20	AAW25147
24	27	33.3	12	21	AAW11981
25	27	33.3	14	14	AAW32256
26	27	33.3	14	16	AAW82648
27	27	33.3	14	21	AAW80806
28	27	33.3	15	16	AAW82650
29	26	32.1	9	15	AAW54036
30	26	32.1	10	19	AAW53901
31	26	32.1	10	21	AAW51015
32	26	32.1	11	21	AAW26012
33	26	32.1	12	21	AAW32250
34	26	32.1	13	21	AAW23579
35	26	32.1	14	16	AAW82649
36	26	32.1	14	19	AAW80529
37	26	32.1	14	22	AAW71412
38	26	32.1	15	10	AAW88158
39	26	32.1	15	10	AAW90817
40	26	32.1	15	17	AAW69888
41	26	32.1	15	18	AAW19442
42	26	32.1	15	18	AAW19453
43	26	32.1	15	18	AAW19453
44	25	30.9	8	8	AAW70132
45	25	30.9	9	8	AAW70130

ALIGNMENTS

RESULT 1	AAW21326	standard; peptide; 9 AA.
ID	AAW21326	
AC	AAW21326	
DF	29-JUL-1997	(first entry)
DE	Glucagon precursor derived signal oligopeptide #31.	
XX	Hydrophilic: signal oligopeptide; hydrophilicity maxima; vaccine; SIV;	
XX	competitive inhibitor; feedback regulator; synthesis; gastrin precursor;	
XX	charge; polarity; farnesyl synthetase; plasminogen activator inhibitor 1;	
XX	hydroxymethylglutaryl coenzyme A reductase; glucagon precursor; rhesus;	
XX	gonadotropin precursor; plasminogen activator inhibitor 2; proteinin;	
XX	gonadotropin amyloid A4; corticotropin releasing factor binding protein;	
XX	apolipoprotein E; herpes virus 1 glycoprotein B; HSV1; human; OMV5;	
XX	herpes virus 2 glycoprotein B; HSV2; collagenase; apolipoprotein A;	
XX	Treponema pallidum membrane protein; TMPA; islet amyloid polypeptide;	
XX	fibroblast MMP1; schistosoma elastase precursor; schistosom; schistosom;	
XX	hepatitis delta antigen; rev protein; HIV; VILV; angiotensinogen.	
OS	Homo sapiens.	
XX	W09519568-A1.	
PN	20-JUL-1995.	
PD	12-JAN-1995;	95WO-US00575.
XX	14-JAN-1994;	94US-0182248.
XX	(RATH/)	RATH M.
XX		

Low T-cell transme
Smooth muscle fibre
Amino acid sequenc
Antigenic decapent
Mycoplasma genital
Amino acid sequenc
Fragment of human
HML-1 alpha-E chain
Anti-fungal and an
Anti-fungal and an
Antibacterial and
Histatin-related p
Linker peptide to
HML-1 alpha-E chain
Peptide which is w
HML-1 alpha-E chain
Anticoagulant pept
Interleukin-1 rece
HHV8 derived human
Human IgE C-epsilo
C-terminal fragmen
C-terminal amino a
HML-1 alpha-E chain
Beta-sheet structu
Peptide forming a
CD66 peptide CD66f
Non-immunogenic am
Human histatin fra
Anti-fungal and an
Anti-fungal and an
Human histatin 3 b
Staphylococcus aur

PI Rath M.
 XX
 DR WPI: 1995-263953/34.
 XX
 PT Identifying signal oligopeptide(s) in protein sequence(s) - shown as
 PT regions of max. hydrophilicity, used in modulating communication
 PT between protein(s)
 PS
 XX
 PS Claim 5; Page 44; 88pp; English.
 CC The sequences given in AAM21201-560 represent hydrophilic signal oligo-
 CC peptides. These signal oligopeptides are localised on the surface
 CC of the protein and are represented by the hydrophilicity maxima of
 CC the protein. These peptides are enriched in charged amino acids
 CC arranged with neutral spacer amino acids. The specific signal
 CC character of these oligopeptides is determined by a characteristic
 CC combination of conformation and charge within the signal sequence.
 CC These oligopeptides may be used as vaccines in the treatment of
 CC human disease, as competitive inhibitors to prevent or reduce the
 CC metabolic action or interaction of a selected protein by blocking
 CC its specific signal sequences, or as therapeutic agents to function
 CC as feedback regulators to reduce synthesis rate of a selected protein.
 CC These peptides may be modified by omitting one or more amino acids at
 CC the N- and/or C-terminal, by substituting one or more amino acids
 CC without consideration of charge and polarity, by substituting one or
 CC more amino acids with amino acid residues with similar charge and/or
 CC polarity, by omitting one or more amino acids or a combination of these.
 CC
 XX
 SQ Sequence 9 AA:
 Query Match 40.7%; Score 33; DB 16; Length 9;
 Best Local Similarity 75.0%; Pred. No. 4.3e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 5 KRHDEQOR 12
 Db 1 krhdefer 8
 RESULT 2
 AAG98093
 ID AAG98093 standard; Peptide; 14 AA.
 AC AAG98093;
 DT 19-SEP-2001 (first entry)
 PT Human SNP associated peptide SEQ ID NO. 735.
 XX
 KW Human; single nucleotide polymorphism; SNP; angiotensin;
 KW 4-hydroxybutyrate; dehydrogenase; protein therapy;
 KW adenosine triphosphate-dependent RNA helicase;
 KW major histocompatibility complex Class I histocompatibility antigen; MHC;
 KW phosphoglycerate kinase; immunosuppressive; immunostimulatory;
 KW antileukemic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
 KW antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine.
 OS
 XX Homo sapiens.
 XX
 PN W0200148245-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 27-DEC-2000; 2000MO-US35346.
 XX
 PR 27-DEC-1999; 99US-0472688.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI: 200A-418297/44.

XX
 PT Polymorphic nucleic acids encoding e.g. angiotensin, dehydrogenase,
 PT adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
 PT kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
 PT diseases and infections -
 PS
 XX
 PS Disclosure: Page 452; 484pp; English.
 CC The invention relates to nucleic acids (AAH79386-AAH80036) encoding
 CC polymorphic variants of proteins (AAG98010-AAG98238) related to
 CC angiotensin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
 CC (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
 CC Class I histocompatibility antigen and/or phosphoglycerate kinase. These
 CC nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
 CC proteins have potential immunosuppressive, immunostimulatory,
 CC antileukemic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
 CC antileukemic, neuroprotective and antimicrobial activity and may be
 CC useful in gene/protein therapy, vaccines, modulation of the expression
 CC and activity of proteins related to angiotensin, 4-hydroxybutyrate,
 CC dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,
 CC major histocompatibility complex (MHC) Class I histocompatibility antigen
 CC and/or phosphoglycerate kinase. Disorders that may be prevented,
 CC diagnosed and/or treated by the above methods include multifactorial
 CC diseases with a genetic component, such as autoimmune diseases (e.g.
 CC rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
 CC erythematosus and Grave's disease), inflammation, cancer (e.g. cancers
 CC of the bladder, brain, breast, colon and kidney, leukemia), diseases of
 CC the nervous system, an infection of pathogenic organisms. They may also
 CC be used to alter phenotypic traits such as longevity, appearance,
 CC strength, speed and endurance.
 CC
 XX
 SQ Sequence 14 AA:
 Query Match 40.7%; Score 33; DB 22; Length 14;
 Best Local Similarity 46.2%; Pred. No. 52;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 1 EIKAKRHDEQORE 13
 Db 1 eqqeqnqqeqq 13
 RESULT 3
 AAM21333
 ID AAM21333 standard; peptide; 11 AA.
 AC AAM21333;
 DT 29-JUL-1997 (first entry)
 PT Glucagon precursor derived signal oligopeptide #38.
 XX
 DE
 KW Hydrophilic; signal oligopeptide; hydrophilicity maxima; vaccine; STV;
 KW competitive inhibitor; feedback regulator; synthesis; gastrin precursor;
 KW charge; polarity; farnesyl synthetase; plasminogen activator inhibitor 1;
 KW hydroxymethylglutaryl coenzyme A reductase; glucagon precursor; rheus;
 KW gonadoliberin precursor; plasminogen activator inhibitor 2; protein;
 KW Alzheimer amyloid A4; corticotropin releasing factor binding protein;
 KW apolipoprotein E; herpes virus 1 glycoprotein B; HSV1; human; OMVVS;
 KW Treponema pallidum membrane protein; TMPA; islet amyloid polypeptide;
 KW fibroblast MMP1; schistosoma elastase precursor; schistosomins;
 KW hepatitis delta antigen; rev protein; HIV; VILV; angiotensinogen.
 OS
 XX Homo sapiens.
 XX
 PN W09519568-A1.
 XX
 PD 20-JUL-1995.
 XX
 PF 12-JAN-1995; 95WO-0500575.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 RN [1]
 RP MEDLINE=84221901; PubMed=6328491;
 RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.J., Jokovic S.,
 RA Rabinowitz M.;
 RT "Characterization of genomic clones specifying rabbit alpha- and beta-
 RT ventricular myosin heavy chains."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).
 DR EMBL; K01698; AAA1415.1;
 KW Myosin.
 FT NON_TER
 SO SEQUENCE 7 AA; 916 MW; 6B1B1AA1E69326B0 CRC64;

Query Match 30.9%; Score 25; DB 6; Length 7;
 Best Local Similarity 66.7%; Pred. No. 4.7e+05;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 RHDEQ 10
 DB 2 KMHDE 7

RESULT 3
 ID P82333 PRELIMINARY; PRT; 14 AA.
 AC P82333;
 DT 01-JUN-2000 (TREMBlrel. 14, Created)
 DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)
 DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF THYLAKOID (SPOT118) (FRAGMENT).
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
 RC STRAIN=CV. DE GRACE; TISSUE=LEAF;
 RX MEDLINE=20181728; PubMed=10715320;
 RA Peltier J.-B., Friso G., Kalume D.E., Roepstorff P., Nilsson F.,
 RA Adamska I., van Wijk K.J.;
 RT "Proteomics of the chloroplast: systematic identification and
 RT targeting analysis of luminal and peripheral thylakoid proteins."
 RT Plant Cell 12:319-341(2000).
 DE 01-SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN OR
 DE PERIPHERY.
 CC 01-DEVELOPMENTAL STAGE: UNFOLDED AND FULLY DEVELOPED LEAVES.
 CC 01-MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.7, ITS MW IS: 28 KDA.
 KW Chloroplast; Thylakoid membrane.
 FT NON_TER
 SO SEQUENCE 14 AA; 1723 MW; CD839374AAF2CA06 CRC64;

Query Match 30.9%; Score 25; DB 10; Length 14;
 Best Local Similarity 44.4%; Pred. No. 2.5e+03;
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 8 DEORELEE 16
 DB 1 EEOEBOEO 9

RESULT 4
 ID 09QCL8 PRELIMINARY; PRT; 13 AA.
 AC 09QCL8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE P40 ORF (FRAGMENT).
 OC Borna disease virus (BDV).
 OS Viruses; ssRNA negative-strand viruses; Mononegavirales.
 OX NCBI_TaxID=12455;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99445641; PubMed=10515835;
 RA Czayan M., Hallensleben W., Hofer M., Pollak S., Sauder C., Bilzer T.,
 RA Blumcke I., Riederer P., Bogerts B., Falkai P., Schwarz M.J.,
 RA Masliah E., Staeheli P., Hufert F.T., Lieb K.;
 RT "Borna disease virus in human brains with a rare form of hippocampal
 RT degeneration but not in brains of patients with common
 RT neuropsychiatric disorders."
 RL J. Infect. Dis. 180:1695-1699(1999).
 DR EMBL; AJ246860; CAB52093.1;
 FT NON_TER
 FT NON_TER
 SO SEQUENCE 13 AA; 1673 MW; 8F80680E534AA0A CRC64;

Query Match 29.6%; Score 24; DB 12; Length 13;
 Best Local Similarity 57.1%; Pred. No. 3.2e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 6 RHDEQOR 12
 DB 1 RHENDR 7

RESULT 5
 ID 09QW8 PRELIMINARY; PRT; 16 AA.
 AC 09QW8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PROTEIN ISOMERASE-RELATED PROTEIN PRECURSOR 71.5 KDA ISOFORM
 DE (FRAGMENT).
 OS Cavia (guinea pigs).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae.
 OX NCBI_TaxID=10140;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96070119; PubMed=8535285;
 RA Bonifacio M.D., Steeves T., Saunders D.M., Sinosich M.J.;
 RT "Isolation of ErbP72 from guinea pig term placenta using heparin
 RT Sepharose affinity chromatography."
 RL Biochem. Mol. Biol. Int. 36:1143-1152(1995).
 SO SEQUENCE 16 AA; 1880 MW; 21B44A5F5767CB17 CRC64;

Query Match 29.6%; Score 24; DB 11; Length 16;
 Best Local Similarity 50.0%; Pred. No. 3.8e+03;
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 8 DEORELEE 15
 DB 3 DEEDDLE 10

RESULT 6
 ID 09TWT4 PRELIMINARY; PRT; 15 AA.
 AC 09TWT4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE HEMOGLOBIN (FRAGMENT).
 OS Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
 OC Lumbricina; Lumbricidae; Lumbricus.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:15:22 ; Search time 10.02 Seconds

(without alignments)
58.547 Million cell updates/sec

Title: US-09-822-110-3

Perfect score: 81

Sequence: 1 EIKAKRHDEQORELEE 16

Scoring table: BLOSUM62

Searched: 100059 seqs, 36664827 residues 780

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	29.6	14	1	RS19_PRUAP
2	24	29.6	16	1	LE06_BIOGL
3	21	25.9	14	1	YMT_P_XANCP
4	20	24.7	15	1	UN01_PINPS
5	19	23.5	9	1	NEUX_HUMAN
6	19	23.5	9	1	NEUX_RAT
7	19	23.5	12	1	RS19_TOBBP
8	19	23.5	13	1	PRC_SALTY
9	18	22.2	10	1	COX4_THOGB
10	17	21.0	10	1	FIBB_CERSI
11	17	21.0	14	1	SCK3_LEIOU
12	16	19.8	5	1	TRM3_ECOLI
13	16	19.8	7	1	UP03_MOUSE
14	16	19.8	12	1	RS19_CLEYE
15	16	19.8	12	1	XYLA_STEVN
16	16	19.8	13	1	NEUT_BUFMA
17	16	19.8	13	1	NEUT_CHICK
18	16	19.8	13	1	NEUT_RANFE
19	16	19.8	13	1	NEUT_TRIVU
20	16	19.8	14	1	DCM_PSECF
21	16	19.8	15	1	ACEM_ACICA
22	16	19.8	16	1	BALI_EUBSP
23	16	19.8	16	1	BL6_VIBRP
24	16	19.8	16	1	YMOR_PSEPU
25	15	18.5	10	1	MALE_KLEPN
26	15	18.5	10	1	SYK_CAMOP
27	15	18.5	13	1	NEUT_CAVPO
28	15	18.5	15	1	GR78_HORSE
29	15	18.5	15	1	UC27_MAIZE
30	15	18.5	15	1	UN04_PINPS
31	15	18.5	16	1	ALLI_CAIYO
32	15	18.5	16	1	LPHI_ECOLI
33	14	17.3	6	1	TRPI_PSEPU

34	14	17.3	8	1	GLUR_HUMAN	P02729	homo sapien
35	14	17.3	8	1	UH09_RAT	P56575	rattus norv
36	14	17.3	10	1	GLEM_HUMAN	P02728	homo sapien
37	14	17.3	10	1	GON3_PETMA	P30948	petromyzon
38	14	17.3	10	1	TRNK_PIG	P01292	scrofa
39	14	17.3	10	1	UH05_RAT	P56573	rattus norv
40	14	17.3	11	1	CS15_BACSU	P81095	baclillus su
41	14	17.3	12	1	LICH_BACLI	P82907	baclillus li
42	14	17.3	12	1	PAZB_VIBBO	P31859	viberra beru
43	14	17.3	12	1	PPK4_PERFU	P82690	periplaneta
44	14	17.3	12	1	R16_GINBI	P36207	ginkgo billo
45	14	17.3	12	1	TAIO_TREME	P01371	trema me

ALIGNMENTS

RESULT 1	RS19_PRUAP	STANDARD:	PRT:	14 AA.
AC	044160;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	30S RIBOSOMAL PROTEIN S19 (FRAGMENT).			
GN	RPSS OR RPSS19.			
OS	Prunus armeniaca phytoplasma.			
OC	Bacteriia: Firmicutes; Bacillus/Clostridium group; Molluscites;			
OC	Actinobacteria: Actinomycetia; Actinobacteria.			
OX	NCBI_TaxID=36389;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94350802; Pubmed=8071198;			
RA	Gundersen D.E., Lee I.M., Rehner S.A., Davis R.E., Kingsbury D.T.;			
RT	"Phylogeny of mycoplasma-like organisms (phytoplasmas): a basis for			
RT	their classification."			
RL	J. Bacteriol. 176:5244-5254(1994).			
CC	-1- FUNCTION: PROTEIN S19 FORMS A COMPLEX WITH S13 THAT BINDS STRONGLY			
CC	TO THE 16S RIBOSOMAL RNA (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.			
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION			
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -			
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS			
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY			
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL			
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/			
CC	OR SEND AN EMAIL TO license@isb-sib.ch).			
DR	EMBL: L26994; AAA83934.1;			
DR	InterPro: IPR002222; RIBOSOMAL_S19.			
DR	PROSITE: PS00323; RIBOSOMAL_S19; PARTIAL.			
KW	Ribosomal protein; rRNA-binding.			
FT	NON_TER			
SQ	SEQUENCE 14 AA: 1756 MW; 4CC029EF8FFFE4A3 CRC64;			
Query Match				
Best Local Similarity 29.6% Score 24; DB 1; Length 14;				
Matches 2; Conservative 20.0%; Pred. No. 4.6e+02;				
Mismatch 7; Indels 1; Gaps 0;				
OY	7 HDEQORELEE 16			
DB	4 HDKDKKTKRK 13			
RESULT 2				
LE06_BIOGL	STANDARD:	PRT:	16 AA.	
AC	P80745;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			

DE HEMOLYMPH 65 KDA LECTIN BG06 (FRAGMENT).
 GN BG06.
 OS Biomphalaria glabrata (Bloodfluke planorb).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
 CC Planorbidae; Biomphalaria.
 CX NCBI_TaxID=6526;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=M-LINE; TISSUE=Hemolymph;
 RX MEDLINE=97385165; PubMed=9238039;
 RA Adema C.M., Hertel L.A., Miller R.D., Loker E.S.;
 RT "A family of fibrinogen-related proteins that precipitates parasite-
 RT derived molecules is produced by an invertebrate after infection."
 CC Proc. Natl. Acad. Sci. U.S.A. 94:8691-8696(1997).
 CC -1- FUNCTION: BINDS AND PRECIPITATES ANTIGENS OF THE PARASITE
 CC ECHINOSTOMA PARASENSI.
 CC -1- SUBCELLULAR LOCATION: SECRETED IN THE HEMOLYMPH.
 CC -1- INDUCTION: BY INFECTION.
 KW Lectin.
 YW NON_TER 1 1
 FT NON_TER 16 16
 FT SEQUENCE 16-AA; 1964 MW; A1665754589EF82C CRC64;
 SQ
 Query Match 29.6%; Score 24; DB 1; Length 16;
 Best Local Similarity 44.4%; Pred. No. 5.2e+02;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 KRHEORE 13
 Db 7 KDNDDOOND 15

RESULT 3
 YWMP_XANCP STANDARD; PRT; 14 AA.
 ID YWMP_XANCP
 AC 003397;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-1993 (Rel. 27, Last annotation update)
 DE HYPOTHETICAL PROTEIN IN FRUK 5' REGION (ORF1) (FRAGMENT).
 GN MTP.
 OS Xanthomonas campestris (pv. campestris).
 CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xanthomonas.
 CX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92011547; PubMed=1655739;
 RA de Crecy-Lagard V., Bouvet O.M., Lejeune P., Danchin A.;
 RT "Fructose catabolism in Xanthomonas campestris pv. campestris.
 RT Sequence of the PTS operon, characterization of the fructose-specific
 RT enzymes."
 CC J. Biol. Chem. 266:18154-18161(1991).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M69242; AAA27600.1;
 DR PIR: C40944; C40944.
 DR Hypothetical protein.
 KW NON_TER 1 1
 FT SEQUENCE 14 AA; 1585 MW; 72EA9AA3CE16CCDB CRC64;
 SO

Query Match 25.9%; Score 21; DB 1; Length 14;
 Best Local Similarity 71.4%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 9 EQORELE 15
 Db 1 EQVRALE 7

RESULT 4
 UNOL_PINPS STANDARD; PRT; 15 AA.
 ID UNOL_PINPS
 AC P81106;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (S1247/S1248) (M150/M151)
 CC (FRAGMENT).
 CC Pinus pinaster (Maritime pine).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 CX NCBI_TaxID=71647;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Needle;
 RA Plomion C., Costa P., Bahrmann N., Frigerio J.M.;
 RT "Genetic analysis of needle proteins in maritime pine. I. Mapping
 RT dominant and codominant protein markers assayed on diploid tissue, in
 RT a haploid-based genetic map."
 CC Silvae Genetica 46:161-165(1997).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Needle;
 RX MEDLINE=99274088; PubMed=10344291;
 RA Costa P., Plomion C., Bauw G., Dudos C., Bahrmann N., Kremer A.,
 RA Frigerio J.M., Plomion C.;
 RT "Separation and characterization of needle and xylem maritime pine
 RT proteins."
 CC Electrophoresis 20:1098-1108(1999).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.5, ITS MW IS: 62 KDA.
 FT NON_TER 1 1
 FT NON_TER 15 15
 FT SEQUENCE 15 AA; 1670 MW; 29CE44CD51E9FCF CRC64;
 SO

Query Match 24.7%; Score 20; DB 1; Length 15;
 Best Local Similarity 75.0%; Pred. No. 1.9e+03;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 HDEQ 10
 Db 1 HDEQ 4

RESULT 5
 NEUX_HUMAN STANDARD; PRT; 9 AA.
 ID NEUX_HUMAN
 AC P04277;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 01-JUL-1989 (Rel. 11, Last annotation update)
 DE NEUROTENSIN-RELATED PEPTIDE (NRP) (KINRENSIN).
 CC Homo sapiens (Human), Bos taurus (Bovine), and
 CC Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606; 9913; 9986;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Human;
 RX MEDLINE=86242180; PubMed=3087352;
 RA Mogard M.H., Kobayashi R., Chen C.F., Lee T.D., Reeve J.R. Jr.,
 RA Shively J.E., Walsh J.H.;
 RT "The amino acid sequence of kinetensin, a novel peptide isolated from
 RT pepsin-treated human plasma: homology with human serum albumin,

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:13:12 ; Search time 12.75 Seconds

(without alignments)
95.592 Million cell updates/sec

Title: US-09-822-110-3

Perfect score: 81

Sequence: 1 EIKAKRHDEQRELEE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:**

1: PIR1:**

2: PIR2:**

3: PIR3:**

4: PIR4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	30.9	7	2	146868
2	25	30.9	14	2	C44823
3	22	27.2	12	2	S47360
4	22	27.2	13	2	A43839
5	21	25.9	14	2	C40944
6	21	25.9	16	2	A35552
7	21	25.9	16	2	F54226
8	20	24.7	11	2	J02307
9	20	24.7	12	2	S16335
10	20	24.7	14	2	E61308
11	20	24.7	14	2	D61308
12	20	24.7	14	2	A39703
13	19	23.5	7	2	A38671
14	19	23.5	8	2	B45800
15	19	23.5	9	2	PM0002
16	19	23.5	11	2	S65395
17	19	23.5	11	2	I52980
18	19	23.5	12	2	T46794
19	19	23.5	13	2	E39778
20	19	23.5	14	2	I54945
21	19	23.5	14	2	B83836
22	19	23.5	15	2	A61247
23	19	23.5	15	2	PH0772
24	19	23.5	15	2	G35141
25	18	22.2	10	2	S77980
26	18	22.2	13	2	A32734
27	18	22.2	14	2	S23639
28	18	22.2	15	2	PN0118
29	18	22.2	15	2	S65717

30	18	22.2	15	2	PA0059	protein QF20021 -
31	18	22.2	15	2	S57584	T cell receptor V-
32	18	22.2	15	2	A53594	calnexin - mouse
33	18	22.2	16	2	C58503	proteoglycan assoc
34	18	22.2	16	2	S28213	glutathione transf
35	17	21.0	12	2	A35585	cytokinin-binding
36	17	21.0	12	2	A34858	proteinase E - bla
37	17	21.0	12	2	B47171	chondroitin sulfat
38	17	21.0	13	2	PC2122	aminotransferase c
39	17	21.0	13	2	J02309	hypothetical 1.6k
40	17	21.0	13	2	J02319	hypothetical 1.6k
41	17	21.0	13	2	PC1149	equinatoxin 1A - s
42	17	21.0	13	2	B47415	mannose-1-phosphat
43	17	21.0	14	2	A48389	leishutoxin III -
44	17	21.0	14	2	PC7079	unidentified 27.2k
45	17	21.0	14	2	S65392	cytochrome-c oxida

ALIGNMENTS

RESULT 1
I46868
alpha-myosin heavy chain ; rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: I46868
R:Friedman, D.J.; Umeda, P.K.; Sluhs, A.M.; Hsu, H.
Proc. Natl. Acad. Sci. U.S.A. 81, 3044-3048, 1984
A:Title: Characterization of genomic clones specifying rabbit alpha- and beta-ventric
A:Reference number: I46868; MUID:84221901
A:Accession: I46868
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7 <FRI>
A:Cross-references: GB:K01698; NID:q165538; PIDN:AAA31415.1; PID:q165539

Query Match 30.9%; Score 25; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.2e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 KRHEQ 10
DB 2 KMHDEE 7
RESULT 2
C44823
synaptosomal-associated protein SNAP-25 peptide 8 - rabbit (fragment)
N:Alternate names: superprotein peptide 8
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 15-Jun-1996
C:Accession: C44823
R:Loewy, A.; Liu, W.S.; Baillinger, C.; Willard, M.B.
J. Neurosci. 11, 3412-3421, 1991
A:Title: The major 35S-methionine-labeled rapidly transported protein (superprotein)
A:Reference number: A44823; MUID:92044785
A:Accession: C44823
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <LOE>
A:Experimental source: visual tissue
A:Note: sequence extracted from NCBI backbone (NCBIP:64253)
C:Keywords: membrane trafficking

Query Match 30.9%; Score 25; DB 2; Length 14;
Best Local Similarity 62.5%; Pred. No. 8.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 8 DEOQRELE 15
DB 2 KMHDEE 7

Db 6 DEORELEE 13

RESULT 3

S47360 T-cell antigen receptor VJ junction beta chain - human

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C:Accession: S47360

R:Lehner, P.J

Submitted to the EMBL Data Library, August 1994

A:Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T

A:Reference number: S47355

A:Accession: S47360

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-12 <LEH>

A:Cross-references: EMBL:Z35684; NID:9527457; PIDN:CA84753.1; PID:9527458

C:Keywords: T-cell receptor

Query Match 27.2%; Score 22; DB 2; Length 12;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 10 DEORELEE 16

Db 2 QREAKE 8

RESULT 4

AA3839 beta-tubulin germ-cell isotype - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C>Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 11-Apr-1995

C:Accession: AA3839

R:Bleker, J.J.; Yazdani-Bulky, M.

Differentiation 50, 15-23, 1992

A:Title: The multiple beta-tubulin genes of Xenopus: isolation and developmental express

A:Reference number: AA3839; MUID:92347627

A:Accession: AA3839

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-15 <BLE>

A:Experimental source: oocyte

A:Note: sequence extracted from NCBI backbone (NCBIP:110252)

Query Match 27.2%; Score 22; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 9 DEORELEE 16

Db 1 EEEBEFE 8

RESULT 5

CA0944 hypothetical protein (1-phosphofructokinase 5' region) - Xanthomonas campestris pv. camp

C:Species: Xanthomonas campestris pv. campestris

C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 18-Jun-1999

C:Accession: CA0944

R:de Crey-Lagard, V.; Bouvet, O.M.M.; Lejeune, P.; Danchin, A.

J. Biol. Chem. 266, 18154-18161, 1991

A:Title: Fructose catabolism in Xanthomonas campestris pv. campestris. Sequence of the F

A:Reference number: AA0944; MUID:92011547

A:Accession: CA0944

A:Molecule type: DNA

A:Residues: 1-14 <DR3>

A:Cross-references: GB:M69242; NID:9155366; PIDN:AAA27600.1; PID:9155367

C:Superfamily: fructose phosphotransferase multiphosphoryltransfer protein: phosphotrans

phosphatase system phosphotransferase-containing protein homology

Query Match 25.9%; Score 21; DB 2; Length 14;
Best Local Similarity 71.4%; Pred. No. 3.1e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 9 DEORELEE 15

Db 1 EQVRALE 7

RESULT 6

A35552 caldesmon - turkey (fragment)

C:Species: Meleagris gallopavo (common turkey)

C>Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 24-Jun-1993

C:Accession: A35552

R:Velaz, L.; Ingraham, R.H.; Chalovich, J.M.

J. Biol. Chem. 265, 2929-2934, 1990

A:Title: Dissociation of the effect of caldesmon on the ATPase activity and on the bi

A:Reference number: A35552; MUID:90153926

A:Accession: A35552

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <VEL>

Query Match 25.9%; Score 21; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 3.6e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 9 DEORELEE 14

Db 4 ERREL 9

RESULT 7

F54226 light-harvesting protein B-830 beta-2 chain - Chromatium purpuratum (fragment)

C:Species: Chromatium purpuratum

C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 05-Jan-1996

C:Accession: F54226

R:Kerfeld, C.A.; Yeates, T.O.; Thorndyke, J.P.

Biochemistry 33, 2178-2184, 1994

A:Title: Purification and characterization of the peripheral antenna of the purple-su

A:Reference number: A54226; MUID:94162224

A:Accession: F54226

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <KER>

A:Keywords: antenna complex; light-harvesting polypeptide

Query Match 25.9%; Score 21; DB 2; Length 16;
Best Local Similarity 44.4%; Pred. No. 3.6e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 8 DEORELEE 16

Db 8 DAQAEFE 16

RESULT 8

JQ2307 hypothetical 1.5K protein - tomato chloroplast (strain Toko)

C:Species: Chloroplast Lycopersicon esculentum (tomato)

C>Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-1995

C:Accession: JQ2307

R:Kawagoe, Y.; Kikuta, Y.

Theor. Appl. Genet. 81, 13-20, 1991

A:Title: Chloroplast DNA evolution in potato (Solanum tuberosum L.).

A:Reference number: JQ2306

A:Accession: JQ2307

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:17:22 ; Search time 12.54 Seconds
(without alignments)
89.726 Million cell updates/sec

Title: US-09-822-110-5
Perfect score: 253
Sequence: 1 PEIKNAKPVSLRLDITLTEAM.....QRELEEEENSDELDISHT 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

total number of hits satisfying chosen parameters: 149863

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_Aa.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTus.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
%score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.5	20.0	30	2	US-08-619-198-7 Sequence 7, Appl
2	50	19.8	35	1	US-08-182-175A-53 Sequence 53, Appl
3	50	19.8	35	1	US-08-474-633A-62 Sequence 62, Appl
4	50	19.8	35	5	PCT-US92-06412-53 Sequence 53, Appl
5	46	18.2	34	2	US-08-690-011A-53 Sequence 53, Appl
6	46	18.2	35	4	US-09-082-279B-72 Sequence 72, Appl
7	45	17.8	28	1	US-08-182-175A-6 Sequence 6, Appl
8	45	17.8	28	5	PCT-US92-06412-6 Sequence 6, Appl
9	45	17.8	35	4	US-09-248-588-26 Sequence 26, Appl
10	44	17.4	20	3	US-08-199-776-21 Sequence 21, Appl
11	44	17.4	20	3	US-08-663-731-21 Sequence 21, Appl
12	44	17.4	20	3	US-08-879-338-21 Sequence 21, Appl
13	44	17.4	20	5	PCT-US95-02044-21 Sequence 21, Appl
14	44	17.4	21	1	US-08-199-776-5 Sequence 5, Appl
15	44	17.4	21	3	US-08-663-731-5 Sequence 5, Appl
16	44	17.4	21	3	US-08-879-338-5 Sequence 5, Appl
17	44	17.4	21	3	US-08-879-338-26 Sequence 26, Appl
18	44	17.4	21	5	PCT-US95-02044-5 Sequence 5, Appl
19	44	17.4	28	4	US-09-082-279B-74 Sequence 74, Appl
20	44	17.4	43	4	US-09-015-030-7 Sequence 7, Appl
21	43.5	17.2	26	2	US-08-690-011A-3 Sequence 3, Appl
22	43	17.0	28	1	US-08-182-175A-49 Sequence 49, Appl
23	43	17.0	28	5	US-08-474-633A-58 Sequence 58, Appl
24	43	17.0	28	5	PCT-US92-06412-49 Sequence 49, Appl
25	43	17.0	35	1	US-08-182-175A-41 Sequence 41, Appl
26	43	17.0	35	1	US-08-474-633A-38 Sequence 38, Appl
27	43	17.0	35	5	PCT-US92-06412-41 Sequence 41, Appl

28	43	17.0	36	4	US-09-082-279B-860	Sequence 860, App
29	43	17.0	33	2	US-08-459-676A-29	Sequence 29, Appl
30	42.5	16.8	43	2	US-08-690-011A-57	Sequence 57, Appl
31	42.5	16.8	49	4	US-08-866-928B-10	Sequence 10, Appl
32	42.5	16.8	49	4	US-08-866-928B-11	Sequence 11, Appl
33	42.5	16.8	49	4	US-08-866-928B-12	Sequence 12, Appl
34	42	16.6	26	2	US-08-523-125-2	Sequence 2, Appl
35	42	16.6	26	3	US-08-660-561A-2	Sequence 2, Appl
36	42	16.6	28	1	US-08-182-175A-4	Sequence 4, Appl
37	42	16.6	28	1	US-08-056-200-13	Sequence 13, Appl
38	42	16.6	28	2	US-08-800-644-13	Sequence 13, Appl
39	42	16.6	28	5	PCT-US92-06412-4	Sequence 4, Appl
40	42	16.6	45	3	US-08-816-346-12	Sequence 12, Appl
41	42	16.6	45	4	US-09-335-411-12	Sequence 12, Appl
42	41.5	16.4	33	2	US-08-690-011A-38	Sequence 38, Appl
43	41	16.2	13	1	US-08-199-776-25	Sequence 25, Appl
44	41	16.2	13	3	US-08-663-731-25	Sequence 25, Appl
45	41	16.2	13	3	US-08-879-338-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-619-198-7
Sequence 7, Application US/08619198
Patent No. 5885831
GENERAL INFORMATION:
APPLICANT: Young, Michael W.
APPLICANT: Senegal, Anita
APPLICANT: Vosschall, Leslie B.
APPLICANT: Price, Jeffrey L.
APPLICANT: Myers, Michael
TITLE OF INVENTION: NUCLEAR LOCALIZATION FACTOR ASSOCIATED
TITLE OF INVENTION: WITH CIRCADIAN RHYTHMS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,198
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-128A CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
US-08-619-198-7

Query Match	20.0%;	Score 50.5;	DB 2;	Length 30;
Best Local Similarity	47.6%;	Pred. No. 6.6;		
Matches 10;	Conservative	6;	Mismatches	4;
			Indels	1;
			Gaps	1;

RESULT 2
US-08-182-175A-53
Sequence 53, Application US/08182175A
Patent No. 5559223
GENERAL INFORMATION:
APPLICANT: Saverio Carl Falco
APPLICANT: Sharon J. Keeler
APPLICANT: Janet A. Rice
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing I
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:

ADDRESS: E. I. du Pont de Nemours and Company
STREET: 1007 Market Street
CITY: Wilmington
STATE: Delaware
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh System, 6.0
SOFTWARE: Microsoft Word, 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/182,175A
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/743,006
FILING DATE: 9 August 1991
ATTORNEY/AGENT INFORMATION:
NAME: Linda Axamethy Floyd
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: BR-1031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 992-4929
TELEFAX: (302) 892-7949
TELEX: 835420
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
IS-08-182-175A-53

Query Match	19.8%	Score 50;	DB 1;	Length 35;
Best Local Similarity	34.6%	Pred. No. 9;		
Matches	9;	Conservative	10;	Mismatches 7;
				Indels 0;
				Gaps 0;
QY	18	EAMEIKAKRHDEOQEELEEEENSD	E 43	
			
db	6	KAMEEKMKMKMEKMKMKMKMKKEE	31	

RESULT 43
US-08-474-633A-62
Sequence 62, Application US/08474633A
Patent No. 5773691
GENERAL INFORMATION:
APPLICANT: E. I. DU PONT DE NEMOURS AND
TITLE OF INVENTION: CHIMERIC GENES AND
APPLICANT: COMPANY

TITLE OF INVENTION: METHODS FOR INCREASING
 TITLE OF INVENTION: INCREASING THE LYSINE
 TITLE OF INVENTION: AND THREONINE CONTENT
 TITLE OF INVENTION: OF THE SEEDS OF PLANTS
 NUMBER OF SEQUENCES: 107
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: E. I. DU PONT DE NEMOURS
 ADDRESSEE: AND COMPANY
 STREET: 1007 MARKET STREET
 CITY: WILMINGTON
 STATE: DELAWARE
 COUNTRY: U.S.A.
 ZIP: 19898
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MICROSOFT WORD VERSION 2.0C
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,633A
 FILING DATE:
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: BARBARA C. STEGELL
 REGISTRATION NUMBER: 30,684
 REFERENCE/DOCKET NUMBER: BB-1037-C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-992-4931
 TELEFAX: 302-773-0164
 TELLEX: 835420
 INFORMATION FOR SEO ID NO: 62:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 35 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-474-633A-62

	Query Match	19.8%	Score 50	DB 1	Length 35
Best Local	Similarity	34.6%	Pred. No. 9		
Matches	9: Conservative	10: Mismatches		7: Indels	0: Gaps
QY	18	EAMEIKARHDEQORELEEEENSD	43		
Db	6	KAMEEKMKMEKMKMEKMKMEK	31		

RESULT 4
PCT-US92-06412-53
Sequence 53, Application PC/TUS9206412
GENERAL INFORMATION:
APPLICANT: Saverio Carl Falco
APPLICANT: Sharon J. Keeler
APPLICANT: Janet A. Rice
TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containin
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. du Pont de Nemours and Company
STREET: 1007 Market Street
CITY: Wilmington
STATE: Delaware
COUNTRY: USA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh System, 6.0
SOFTWARE: Microsoft Word, 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06412
FILING DATE: 19920807
CLASSIFICATION: 530

[illegible]

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	24.5	49	22	AA021894	Peptide #8328 <i>encc</i>
2	62	24.5	49	22	AA038220	Peptide #12257 <i>encc</i>
3	61	24.1	37	18	AA045249	Peptide <i>conjugate</i>
4	56	22.1	41	18	AA006736	Calreticulin C-dom
5	54	21.3	31	22	AA033848	Peptide #7685 <i>encc</i>
6	54	21.3	41	22	AA016668	Peptide #3102 <i>encc</i>
7	54	21.3	41	22	AA029152	Peptide #3189 <i>encc</i>
8	54	21.3	41	22	AA004384	Peptide #3066 <i>encc</i>
9	54	21.3	50	22	AA015044	Peptide #1478 <i>encc</i>
10	54	21.3	50	22	AA027486	Peptide #1523 <i>encc</i>
11	54	21.3	50	22	AA002775	Peptide #1457 <i>encc</i>

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:18:12 ; Search time 22.07 Seconds
(without alignments)
331.383 Million cell updates/sec

Title: US-09-822-110-5
Perfect score: 253
Sequence: 1 PFIKNAKPVSIIRLDITTEAM.....ORELEEEENSDDELDHSMT 50

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 27079

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	22.1	41	5	09TX57
2	52	20.6	45	5	09VFU5
3	49.5	19.6	50	11	096219
4	49	19.4	41	10	096414
5	49	19.4	44	12	069202
6	48.5	19.2	44	12	069199
7	48	19.0	44	10	003396
8	48	18.8	44	5	09VR85
9	47.5	18.8	44	4	016264
10	47	18.6	44	12	069203
11	47	18.6	44	12	069192
12	47	18.6	44	12	069198
13	46.5	18.6	46	5	009363
14	46.5	18.2	50	10	09C6E9
15	46	18.2	44	5	09N7M4
16	44	17.4	44	12	069195
17	43	17.0	37	11	062620
18	43	17.0	48	4	012978
19	43	17.0	50	2	006970

20	43	17.0	50	12	09JUE6	09JUE6 vaccinia vi
21	42	16.6	43	10	09M5U2	09M5U2 nicotiana t
22	42	16.6	44	12	0993M8	0993M8 human immu
23	42	16.6	45	2	09KED2	09KED2 bacillus ha
24	42	16.6	46	5	020835	020835 caenorhabdi
25	41.5	16.4	45	10	09M5U1	09M5U1 nicotiana t
26	41.5	16.4	50	2	099725	099725 staphylococ
27	41	16.2	44	5	09V7I6	09V7I6 drosophila
28	41	16.2	49	13	090340	090340 cyprinus ca
29	41	16.2	35	4	09UE13	09UE13 homo sapien
30	40.5	16.0	35	13	090347	090347 coturnix co
31	40.5	16.0	43	10	09LUZ2	09LUZ2 arabidopsis
32	40.5	16.0	50	10	09LUX4	09LUX4 human para
33	40	15.8	47	12	09BDE6	09BDE6 bos taurus
34	40	15.8	49	6	09W86	09W86 drosophila
35	40	15.8	50	12	064879	064879 unidentified
36	40	15.8	50	12	09MX3	09MX3 human para
37	39	15.4	16	11	0900W8	0900W8 cavia (guin
38	39	15.4	20	4	015969	015969 homo sapien
39	39	15.4	44	11	060500	060500 cricetus
40	39	15.4	44	12	P87665	P87665 duck adenov
41	39	15.4	48	5	09W86	09W86 drosophila
42	39	15.4	50	4	09P0K4	09P0K4 homo sapien
43	38	15.0	23	6	029394	029394 canis fami
44	38	15.0	34	2	069770	069770 pseudomonas
45	38	15.0	48	2	09ZL87	09ZL87 helicobacte

ALIGNMENTS

RESULT 1
ID 09TX57 PRELIMINARY: PRT: 41 AA.
AC 09TX57:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CADESMON HOMOLOG (FRAGMENT).
OS Mytilus.
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
OC Mytiloidae; Mytilidae.
OX NCBI_TaxID=6548;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265000; PubMed=8205392;
RA Cizmadi A.M., Bonet-Kerrache A., Nyitray L., Mornet D.;
RT "Purification and Properties of caldesmon-like protein from molluscan
smooth muscle."
RL Comp. Biochem. Physiol. 108B:59-63(1994).
FT NON_TER 1 1
FT NON_TER 41 41
FT NON_TER 51 51
SO SEQUENCE 41 AA; 5512 MW; DD511D27BA04A555 CRC64;

Query Match 22.1%; Score 56; DB 5; Length 41;
Best Local Similarity 41.7%; Pred. No. 47;
Matches 10; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
OY 21 EIKAKRHDEQORELEEEENSD 44
DB 18 EDRRREDEERRRREERKREE 41
RESULT 2
ID 09VFU5 PRELIMINARY: PRT: 45 AA.
AC 09VFU5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG14371 PROTEIN.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga C.P., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu I., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Botkin P.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrel J., Fleischmann W.,
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelfand J.H., Glasser K.,
 RA Glisler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiner K., Remington K., Sanders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003700; AAF54954.1;
 DR FlyBase: FBgn0040555; CG14371.
 SR
 SQ SEQUENCE 45 AA; 5235 MW; 93A8CBA8E0F54082 CRC64;

Query Match 20.6%; Score 52; DB 5; Length 45;
 Best Local Similarity 27.8%; Pred. No. 1.3e+02;
 Matches 10; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

OY 10 SLIRDLITAMEIKAKRHDEQRLREFEENSDEDE 45
 ID 062619 PRELIMINARY; PRT; 50 AA.
 AC 062619;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE FETAL TROPONIN T 3 (FRAGMENT).
 OC Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE, SKELETAL;
 RX MEDLINE=93345743; PubMed=83444466;
 RA Briggs M.M., Schachar F.,
 RT "Origin of fetal tropoin T: developmentally regulated splicing of a
 RT new exon in the fast tropoin T gene."
 RL Dev. Biol. 158:503-509(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE, SKELETAL;
 RA Briggs M.M.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U04980; AAA16033.1;
 KW Alternative splicing
 FT NON_TER 50
 SQ SEQUENCE 50 AA; 5982 MW; A4ACC3066914925D CRC64;

Query Match 19.4%; Score 49; DB 10; Length 50;
 Best Local Similarity 36.1%; Pred. No. 2.5e+02;
 Matches 13; Conservative 5; Mismatches 11; Indels 7; Gaps 1;

OY 17 TEAMEIKAKRHDEQRLREFEENSDEDE 45
 ID 096414 PRELIMINARY; PRT; 41 AA.
 AC 096414;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE BETA-TUBULIN 3 (FRAGMENT).
 GN TUBB.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OC NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, KINTOKI;
 RA Okamura S., Sonehara K., Naito K., Ohkawa H., Kuramori S., Tatsuta M.,
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U64430; AAA7936.1;
 DR Mendel: 13640; Dauga; TUBB; 13640.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 KW GTP-binding.
 FT NON_TER 1
 SQ SEQUENCE 41 AA; 4945 MW; A5236AAE45F10274 CRC64;

Query Match 19.4%; Score 49; DB 10; Length 41;
 Best Local Similarity 32.4%; Pred. No. 2.4e+02;
 Matches 11; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

OY 10 SLIRDLITAMEIKAKRHDEQRLREFEENSDEDE 43
 ID 069202 PRELIMINARY; PRT; 44 AA.
 AC 069202;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

OC Termopsidae; zoothermopsids.
 OX NCBI_TaxID=136037;
 RN [1]
 RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND AMIDATION.
 RX TISSUE=Head;
 RC MEDLINE=21107389; PubMed=11179807;
 RA Baldwin D.C., Schlegel K.M., Furuya K., Lehmberg E., Schooley D.A.;
 RT "Isolation and identification of a diuretic hormone from zoothermopsids neandensis."
 RL Peptides 22:147-152(2001).
 CC -1- FUNCTION: REGULATION OF FLUID SECRETION. STIMULATES PRIMARY URINE SECRETION BY MALPIGHIAN TUBULES AND CAUSES A DOSE-DEPENDENT STIMULATION OF CAMP LEVELS IN THE TUBULES.
 CC -1- MASS SPECTROMETRY: MW=5329.0; METHOD=MALDI.
 CC -1- SIMILARITY: BELONGS TO THE SAUVAGINE/CORTICOTROPIN-RELEASING FACTOR/UROTENSIN I FAMILY OF PEPTIDES.
 CC Interpro: IPR00187; CRF.
 DR Interpro: IPR003621; Diuretic_hormn.
 DR Pfam: PF00473; CRF; 1.
 DR Prodom: PD014750; Diuretic_hormn; 1.
 DR SMART: SM00039; CRF; 1.
 DR PROSITE: PS00511; CRF; 1.
 DR Hormone: Amidation.
 KM MOD_RES 46
 FT MOD_RES 46
 SQ SEQUENCE 46 AA; 5329 MW; 6938C02E11F151A5 CRC64;

Query Match 15.4%; Score 39; DB 1; Length 46;
 Best Local Similarity 23.1%; Pred. No. 5.6e+02;
 Matches 9; Conservative 10; Mismatches 18; Indels 2; Gaps 1;
 OY 1 PFKNKAPVSLRDITLMEAIKRRHDEQRELEEE 39
 DB 5 PSLSTVNPDLVLRQL--LLEIARRRMRSQDQIOANRE 41

RESULT 3
 PS11_PSEFL STANDARD; PRT; 39 AA.
 AC P80694;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PHOSPHATE STARVATION INDUCIBLE PROTEIN 1 (PS11) (FRAGMENT).
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=294;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=AG1;
 RX MEDLINE=97237729; PubMed=9084184;
 RA Leopold K., Jacobsen S., Nybroe O.;
 RT "A phosphate-starvation-inducible outer-membrane protein of Pseudomonas fluorescens Ag1 as an immunological phosphate starvation marker."
 RL Microbiology 143:1019-1027(1997).
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE.
 KW Outer membrane.
 FT NON_TER 39
 SQ SEQUENCE 39 AA; 4290 MW; 02280E1131254ACF-CRC64;

Query Match 15.0%; Score 38; DB 1; Length 39;
 Best Local Similarity 28.6%; Pred. No. 5.9e+02;
 Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

OY 12 LRDLTEAMEIKRRHDEQRELEEE 39
 DB 6 LIDMLKANGQISASQYTELOAEIANKOK 33

RESULT 4

DIUH_MUSDO STANDARD; PRT; 44 AA.
 AC P41537;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE DIURETIC HORMONE (DH) (DIURETIC PEPTIDE) (DP).
 OS Musca domestica (House fly), and stomoxys calcitrans (Stable fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pharygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Muscoidea; Muscidae; Musca.
 OX NCBI_TaxID=7370, 35570;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95083479; PubMed=7991460;
 RA Clotiens F.L., Holman G.M., Coast G.M., Totty N.F., Hayes T.K., Kay I.I., Maillet A.I., Wright M.S., Chung J.-S., Truong O., Bull D.L.;
 RT "Isolation and characterization of a diuretic peptide common to the house fly and stable fly."
 RL Peptides 15:971-979(1994).
 CC -1- FUNCTION: REGULATION OF FLUID SECRETION. STIMULATES PRIMARY URINE SECRETION BY MALPIGHIAN TUBULES AND CAUSES A DOSE-DEPENDENT STIMULATION OF CAMP LEVELS IN THE TUBULES. MAY ACT AS CLEARANCE PEPTIDE IN THAT IT MAY REMOVE METABOLIC WASTE FROM THE HEMOLYMPH.
 CC -1- SIMILARITY: BELONGS TO THE SAUVAGINE/CORTICOTROPIN-RELEASING FACTOR/UROTENSIN I FAMILY OF PEPTIDES.
 CC Interpro: IPR00187; CRF.
 DR Interpro: IPR00473; CRF; 1.
 DR Pfam: PF00473; CRF; 1.
 DR SMART: SM00039; CRF; 1.
 DR PROSITE: PS00511; CRF; 1.
 DR Hormone: Amidation.
 KM MOD_RES 44
 FT MOD_RES 44
 SQ SEQUENCE 44 AA; 5181 MW; 67628086D21487B6 CRC64;

Query Match 15.0%; Score 38; DB 1; Length 44;
 Best Local Similarity 25.7%; Pred. No. 6.6e+02;
 Matches 9; Conservative 9; Mismatches 15; Indels 2; Gaps 1;
 OY 1 PFKNKAPVSLRDITLMEAIKRRHDEQRELE 35
 DB 3 PSLSTVNPDLVLRQL--LLEIARRRMKNENRNOVE 35

RESULT 5
 OSTC_XENLA STANDARD; PRT; 49 AA.
 AC P40147;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE OSTEOCALCIN (GAMMA-CARBOXYGLUTAMIC ACID-CONTAINING PROTEIN) (BONE GLA-PROTEIN) (BGP).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96139691; PubMed=8567186;
 RA Cancelli M., Williamson M.K., Price P.A.;
 RT "Amino-acid sequence of bone gla protein from the African clawed toad Xenopus laevis and the fish Sparus aurata."
 RL Int. J. Pept. Protein Res. 46:419-423(1995).
 CC -1- FUNCTION: THIS BONE PROTEIN, CONSTITUTES 1-2% OF THE TOTAL PROTEIN OF BONE. IT BINDS STRONGLY TO APATITE AND CALCIUM.
 CC -1- PTH: GAMMA-CARBOXYGLUTAMIC ACID RESIDUES ARE FORMED BY VITAMIN K DEPENDENT CARBOXYLATION. THESE RESIDUES ARE ESSENTIAL FOR THE BINDING OF CALCIUM.
 CC -1- SIMILARITY: BELONGS TO THE OSTEOCALCIN / MATRIX GLA-PROTEIN FAMILY.
 DR Interpro: IPR002384; GLA_bone.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:17:42 ; Search time 13.27 Seconds
(without alignments)
287.018 Million cell updates/sec

Title: US-09-822-110-5

Perfect score: 253
Sequence: 1 PIRKAKPVSLRDLTEAM.....QRELEEEENSDDELSDHT 50

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 11081

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: PIR-68:***
1: pirl:***
2: pirl2:***
3: pirl3:***
4: pirl4:***

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	19.8	49	2	A56388
2	49	19.4	41	2	T14350
3	47.5	18.8	44	2	I51889
4	47	18.6	39	2	P00011
5	46.5	18.4	50	2	F86472
6	45	17.8	44	2	C34327
7	44	17.4	38	2	I66797
8	44	17.4	43	2	S02031
9	43	17.0	48	2	S74261
10	42	16.6	45	2	A83765
11	42	16.6	46	2	T16060
12	40	15.8	40	2	A60645
13	39	15.4	44	2	I48118
14	39	15.4	46	2	A48542
15	38	15.0	25	2	S72229
16	38	15.0	36	2	JA0173
17	38	15.0	42	2	F43259
18	38	15.0	48	2	E71901
19	37.5	14.8	48	2	S61469
20	37.5	14.8	48	2	S61472
21	37	14.6	42	2	F82049
22	37	14.6	44	2	A25006
23	37	14.6	48	2	D64614
24	36.5	14.4	48	2	T07305
25	36	14.2	41	2	T48342
26	36	14.2	42	2	A39124
27	36	14.2	48	2	I46522
28	36	14.2	49	2	S08452
29	35.5	14.0	48	2	E85646

30	35.5	14.0	49	2	J01200	hypothetical 5.4k
31	35	13.8	31	2	G81558	hypothetical prote
32	35	13.8	32	2	I46523	tropomyosin T 4f - ra
33	35	13.8	34	2	S56118	myosin II heavy ch
34	35	13.8	42	2	I65746	tropomyosin - huma
35	35	13.8	42	2	T07285	hypothetical prote
36	35	13.8	45	2	T11960	ribosomal protein
37	35	13.8	48	2	S02208	osteocalcin - emu
38	35	13.8	49	2	H83820	hypothetical prote
39	35	13.8	50	2	T06541	probable NADPH--fe
40	35	13.8	50	2	S77712	mitochondrial oute
41	34	13.4	23	2	B04348	internal peptide V
42	34	13.4	23	2	A04348	internal peptide V
43	34	13.4	39	2	G82287	hypothetical prote
44	34	13.4	39	2	C83904	hypothetical prote
45	34	13.4	43	2	S41388	protein 3a - human

ALIGNMENTS

RESULT 1
A56388
dopamine- and cAMP-regulated neuronal phosphoprotein DARPP-32 - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: A56388
R:Desdoutils, F.; Cohen, D.; Nairn, A.C.; Greengard, P.; Gilraut, J.A.
J. Biol. Chem. 270, 8772-8778, 1995
A:Title: Phosphorylation of DARPP-32, a dopamine- and cAMP-regulated phosphoprotein,
A:Reference number: A56388; MUID:95236371
A:Accession: A56388
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-49 <DE>
C:Keywords: phosphoprotein
F:7.35.43/Binding site: phosphate (Ser) (covalent) (by casein kinase I) #status exper

Query Match 19.8%; Score 50; DB 2; Length 49;
Best Local Similarity 37.5%; Pred. No. 1e+02;
Matches 9; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Oy 21 EIKAKRHDEQRELEEEENSDSD 44
Db 11 ELGYRQDEDEDEDEDEDEDEED 34

RESULT 2

T14350
tubulin beta-3 chain - carrot (fragment)
C:Species: Daucus carota (carrot)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000
C:Accession: T14350
R:Okamura, S.; Somehara, K.; Naito, K.; Okawa, H.; Kuramori, S.; Tatsuta, M.; Minami
submitted to the EMBL Data Library, July 1996
A:Description: Characterization of beta-tubulin genes of carrot.
A:Reference number: Z17999
A:Accession: T14350
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-41 <OK>
A:Cross-references: EMBL:U64430; NID:g1553124; RID:g1553125
A:Experimental source: cultivar Kintoki
C:Comment: Tubulin is a dimer of alpha and beta chains and is found in the microtubul
changeable site on its beta chain and at a nonexchangeable site not yet identified. T
C:Complex: heterodimer: alpha and beta chain
C:Superfamily: tubulin
C:Keywords: heterodimer

Query Match 19.4%; Score 49; DB 2; Length 41;
Best Local Similarity 32.4%; Pred. No. 1.1e+02;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

C; Keywords: skeletal muscle

13 LBDITTEAMETKAKRBHDEOORELEEEENDEDEL 46
Ov12 LRDLTEAMEIKAKRHDEQORELEEEENSDDEL 46
QY

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:05:01 ; Search time 19.84 Seconds
(without alignments)
907.381 Million cell updates/sec

Title: US-09-822-110-2

Perfect score: 2564
Sequence: 1 MEGPAPAKSKLKKLSLSDSLT.....AKROPILDAMDAKKRRQNF 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

al number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2564	100.0	491	1	STK3_HUMAN	Q13188 homo sapien
2	1968.5	76.8	487	1	STK4_HUMAN	Q13043 homo sapien
3	817	31.9	443	1	ST24_HUMAN	O9Y660 homo sapien
4	796	31.0	426	1	ST25_HUMAN	O00506 homo sapien
5	792	30.9	426	1	ST25_MOUSE	O97241 mus musculu
6	697.5	27.2	968	1	ST10_HUMAN	O94804 homo sapien
7	693	27.0	966	1	ST10_MOUSE	O35098 mus musculu
8	634	24.7	544	1	PAK1_RAT	P33465 rattus norv
9	633	24.7	545	1	PAK1_HUMAN	O75914 homo sapien
10	630	24.6	544	1	PAK3_HUMAN	O62829 rattus norv
11	630	24.6	544	1	PAK3_MOUSE	O61036 mus musculu
12	629	24.5	544	1	PAK3_MOUSE	O29502 oryctolagus
13	628.5	24.5	524	1	PAK2_RABIT	O64303 rattus norv
14	628.5	24.5	524	1	PAK2_RAT	O64303 rattus norv
15	626.5	24.4	524	1	PAK1_HUMAN	O13177 homo sapien
16	626	24.4	545	1	PAK1_MOUSE	O88643 mus musculu
17	618.5	24.1	982	1	SIL1L_CAEEL	P46549 caenorhabdi
18	595.5	23.2	658	1	PAK1_SCHRO	P50527 schizosacch
19	593	23.1	1080	1	MIL5_CAEEL	O23356 caenorhabdi
20	592	23.1	490	1	SPS1_YEAST	P08458 saccharomyc
21	572	22.3	971	1	CLAD_CANAL	O14427 candida alb
22	564.5	22.0	939	1	ST20_YEAST	O04297 saccharomyc
23	554	21.6	1230	1	ST20_CANAL	O92212 candida alb
24	552	21.5	589	1	SHK2_SCHRO	O10056 schizosacch
25	552	21.5	1080	1	NK1L_YEAST	P36692 saccharomyc
26	548.5	21.4	1135	1	NINS_DROME	P10677 drosophila
27	548.5	21.4	1501	1	NINL_DROME	P10676 drosophila
28	533.5	20.8	547	1	SPAK_HUMAN	O94968 homo sapien
29	530.5	20.8	556	1	SPAK_MOUSE	O88506 mus musculu
30	530.5	20.7	553	1	SPAK_RAT	O98156 rattus norv
31	528.5	20.5	842	1	CLA4_YEAST	P48562 saccharomyc
32	520	20.3	591	1	PAK4_HUMAN	O96013 homo sapien
33	516.5	20.1	1062	1	CC7_SCHPO	P41892 schizosacch

34	515	20.1	719	1	PAK5_HUMAN	O9P286 homo sapien
35	509.5	19.9	655	1	SKM1_YEAST	O12469 saccharomyc
36	501	19.5	974	1	CC15_YEAST	P27636 saccharomyc
37	490	19.1	392	1	MPK1_HUMAN	O02750 homo sapien
38	490	19.1	392	1	MPK1_MOUSE	P31938 mus musculu
39	490	19.1	392	1	MPK1_RABIT	P28678 oryctolagus
40	490	19.1	392	1	MPK1_RAT	O01986 rattus norv
41	489	19.1	393	1	MPK1_CRICR	O63980 cricetus
42	486	19.0	388	1	MPK1_SERCA	O91447 serinus can
43	479	18.7	394	1	MPK1_XENLA	O05116 xenopus lae
44	478.5	18.7	401	1	MPK2_MOUSE	O63932 mus musculu
45	477	18.6	400	1	MPK2_RAT	P36506 rattus norv

ALIGNMENTS

RESULT	ID	STK3_HUMAN	STANDARD:	PRT:	491 AA.
AC	Q13188	Q15801: Q15445;			
DT	20-AUG-2001	(Rel. 40, Created)			
DT	20-AUG-2001	(Rel. 40, Last sequence update)			
DE	SERINE/THREONINE PROTEIN KINASE 3 (EC 2.7.1.37) (STE20-LIKE KINASE				
DE	MST2) (MST-2) (MAMMALIAN STE20-LIKE PROTEIN KINASE 2)				
DE	(SERINE/THREONINE PROTEIN KINASE KRS-1).				
GN	STK3 OR MST2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
NC	NCBI_TaxID=9606;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96144292; PubMed=8566796;				
RA	Creasy C.L., Chernoff J.;				
RT	"Cloning and characterization of a member of the MST subfamily of				
RT	Ste20-like kinases."				
RL	Gene 1671303=306(1995).				
RM	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96413604; PubMed=8816758;				
RA	Taylor L.K., Wang H.C., Erikson R.L.;				
RT	"Newly identified stress-responsive protein kinases, Krs-1 and Krs-				
RT	2."				
RL	[3]				
RP	Proc. Natl. Acad. Sci. U.S.A. 93:10999-10104(1996).				
RX	SEQUENCE OF 96-203 FROM N.A.				
RX	MEDLINE=94100173; PubMed=8274451;				
RA	Schultz S.J., Nigg E.A.;				
RT	"Identification of 21 novel human protein kinases, including 3 members				
RT	of a family related to the cell cycle regulator NIMA of Aspergillus				
RT	nidulans."				
RL	Cell Growth Differ. 4:821-830(1993).				
CC	FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT				
CC	MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY				
CC	SIMILARITY).				
CC	- CATALYTIC ACTIVITY: ATP + A PROTEIN - ADP + A PHOSPHOPROTEIN.				
CC	- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).				
CC	- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,				
CC	SKLETERAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT				
CC	HEART, LUNG AND BRAIN TISSUES.				
CC	- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.				
CC	STE20 SUBFAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				

```
Query Match      100.0%; Score 2564; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 4, 5e-142;
Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

RESULT	2
STK4_HUMAN	
ID	STK4_HUMAN
AC	Q13043; Q15802; Q9NT24;

DT	20-AUG-2001	(Rel. 40. Created)	
DT	20-AUG-2001	(Rel. 40. Last sequence update)	
DT	20-AUG-2001	(Rel. 40. Last annotation update)	
DE	SERINE/THREONINE PROTEIN KINASE 4 (EC 2.7.1.37) (STE20-LIKE KINASE		
DE	MS1) (MST-1) (MAMMALIAN STE20-LIKE PROTEIN KINASE 1)		
DE	(SERINE/THREONINE PROTEIN KINASE KRS-2).		
GN	STE4 OR MST1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95394929; PubMed=7665586;		
RT	Creasy C.L., Chernoff J.;		
RT	"Cloning and characterization of a human protein kinase with homology		
RT	to ste20.";		
RT	J. Biol. Chem. 270:21695-21700(1995).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96413604; PubMed=8816758;		
RT	Taylor L.K., Wang H.C., Erikson R.L.;		
RT	"Newly identified stress-responsive protein kinases, Krs-1 and Krs-		
RT	2.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).		
RN	[3]		
RP	SEQUENCE OF 1-435 FROM N.A.		
RA	Laird G.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.		
CC	-1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS		
CC	THE KINASE ACTIVITY.		
CC	-1- SUBCELLULAR LOCATION: CYTOSOLASMIC (BY SIMILARITY).		
CC	-1- TISSUE SPECIFICITY: UNIDUOUSLY EXPRESSED.		
CC	-1- PH: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.		
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
CC	STE20 SUBFAMILY.		
CC	-----		
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION		
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -		
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS		
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY		
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL		
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/		
CC	OR SEND AN EMAIL TO license@isb-sib.ch).		
CC	-----		
CC	EMBL; U18297; AAA83254.1; -		
CC	EMBL; U60207; AAB17262.1; -		
CC	EMBL; AL109839; CAB89421.1; -		
CC	HSSP; P24941; 1HCL.		
CC	MIM; 604965; -		
CC	InterPro: IPR000719; Euk_pkinase.		
CC	InterPro: IPR002290; Ser_thr_kin_actsite.		
CC	InterPro: IPR001245; Tyr_kin.		
CC	Pfam: PF00069; pkinase; 1.		
CC	PRINTS: PR00109; TYRKINASE.		
CC	SMART: SMO0220; S_TKC; 1.		
CC	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.		
CC	PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.		
CC	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
CC	Transferrase; Serine/threonine protein kinase; ATP-binding.		
CC	DOMAIN	30	281
CC	NE_BIND	36	44
CC	BINDING	59	59
CC	ACT_SITE	149	149
CC	DOMAIN	373	378
CC	CONFLICT	222	222
CC	CONFLICT	312	312
CC	SEQUENCE	487 AA; 55630 MW; 1507585BEC5F77D5C CRC64;	

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: February 15, 2002, 07:04:32 ; Search time 53.55 Seconds
(without alignments)
1341.171 Million cell updates/sec

Title: US-09-822-110-2
Perfect score: 2564
Sequence: 1 MEOPAPPSKRLKLSLSEDLT.....AKROPILDMADAKKROQNF 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues
al number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2462	96.0	491	11 054748	054748.rattus norv
2	2442	95.2	497	11 09J110	09J110 mus musculu
3	2110	82.3	445	11 060877	060877 mus musculu
4	1961.5	76.5	487	11 09J111	09J111 mus musculu
5	1555	60.6	321	11 09CW82	09CW82 mus musculu
6	1460.5	57.0	478	5 062571	062571 suberites d
7	1411.5	55.1	596	5 09WB4	09WB4 drosophila
8	1315.5	51.3	497	5 09NB31	09NB31 caenorhabdi
9	1076.5	42.0	461	5 061125	061125 dictyosteli
10	816.5	31.8	642	5 09VEK3	09VEK3 drosophila
11	815.5	31.8	416	4 09BXC4	09BXC4 homo sapien
12	815.5	31.8	416	11 09J9T2	09J9T2 mus musculu
13	813	31.7	416	4 09P289	09P289 homo sapien
14	811.5	31.6	431	11 099KH8	099KH8 mus musculu
15	802.5	31.3	478	5 061122	061122 dictyosteli
16	800	31.2	653	5 022553	022553 caenorhabdi
17	775	30.2	1120	10 09LOA1	09LOA1 arabidopsis
18	774.5	30.2	836	10 024527	024527 arabidopsis
19	772.5	30.1	842	10 09FNU3	09FNU3 oryza sativ

20	765.5	29.9	825	10 09ARL7	09ARL7 hordeum vul
21	714.5	27.9	1202	11 09W041	09W041 mus musculu
22	714.5	27.9	1233	11 054988	054988 mus musculu
23	712	27.8	690	10 09LDM6	09LDM6 arabidopsis
24	706	27.5	676	10 09ZPD6	09ZPD6 brassica na
25	705.5	27.5	1615	4 09N58	09N58 homo sapien
26	701.5	27.4	1218	5 09V8R6	09V8R6 drosophila
27	701.5	27.4	1231	11 055092	055092 cavia porce
28	701.5	27.4	1297	4 09UKD9	09UKD9 homo sapien
29	701.5	27.4	1305	4 09UKR3	09UKR3 homo sapien
30	701.5	27.4	1352	4 09UKR2	09UKR2 homo sapien
31	701.5	27.4	1360	4 09UKR5	09UKR5 homo sapien
32	700.5	27.3	1300	11 09JUM92	09JUM92 mus musculu
33	700.5	27.3	1268	11 09JUM52	09JUM52 mus musculu
34	698	27.2	1276	4 09UKR1	09UKR1 homo sapien
35	698	27.2	1323	4 09UKR0	09UKR0 homo sapien
36	698	27.2	1331	4 09UKR4	09UKR4 homo sapien
37	697.5	27.2	1152	4 092603	092603 homo sapien
38	697.5	27.2	1204	4 000211	000211 homo sapien
39	697.5	27.2	1235	4 09H262	09H262 homo sapien
40	697.5	27.2	471	3 014305	014305 schizosacch
41	697	27.2	1206	11 008815	008815 rattus norv
42	696	27.1	1295	4 09P2R8	09P2R8 homo sapien
43	694.5	27.1	1303	4 09PIX1	09PIX1 homo sapien
44	694.5	27.1	1303	4 09PIX1	09PIX1 homo sapien
45	693	27.0	1165	4 095819	095819 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	491 AA.
ID	054748			
AC	054748			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	MSR2 KINASE.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Aurisicchio, L., DiLauro R., Zannini M.;			
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ001529; CAA04814.1;			
DR	HSSP; P00518; 1PHK			
DR	InterPro: IPR000719; Euk PKinase.			
DR	InterPro: IPR002290; Ser thr_kin_acsite.			
DR	Pfam; PF00069; PKinase; 1.			
DR	SMART; SM00220; S_TRC; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP_1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
KW	ATP-binding; Transferase			
SG	SEQUENCE 491 AA; 56121 MW; 90FA6B020E7FFD62 CRC64;			
Query Match	96.0%; Score 2462; DB 11; Length 491;			
Best Local Similarity	96.1%; Pred. No. 5.6e-160;			
Matches	472; Conservative 9; Mismatches 10; Indels 0; Gaps 0;			
QY	1 MEOPAPPSKRLKLSLSEDLTROPPEVFLVLEKLGSGSVFKAIHKSGGVAAIKQVPV 60			
DB	1 MEOPAPPSKRLKLSLSEDLTROPPEVFLVLEKLGSGSVFKAIHKSGGVAAIKQVPV 60			
QY	61 ESDLOEIKKEISIMQCCSPYVVKYGYFKNTDLMIYMEYCGAGSVDIIRLRKLTLE 120			
DB	61 ESDVOEIKKEISIMQCCSPYVVKYGYFKNTDLMIYMEYCGAGSVDIIRLRKLTLE 120			
QY	121 DEATILIKSTLGLLEYLFHMRKIRHIDKAGNILLTEGHAKLADPGVAGOLTDTPMAKRN 180			

```

Db 121 DEATILKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGNAKLADPGVAGQLTDTMAKRN 180
Qy 181 VIGTFPMMAPEVIOEIGYNCVADIMSLGITSIEMAEKGPYPADIPHMAIEMIPNPPPT 240
Db 181 VIGTFPMMAPEVIOEIGYNCVADIMSLGITSIEMAEKGPYPADIPHMAIEMIPNPPPT 240
Qy 241 FRKPELMSDDFTDFPKKCLVKNPEORATATOLLHPFIKNKPPVSLRLDLTEAMEIKAK 300
Db 241 FRKPELMSDDFTDFPKKCLVKNPEORATATOLLHPFIKNKPPVSLRLDLTEAMEIKAK 300
Qy 301 RHDEOORELEEEENEDEDELDSHTMVKTSVEGCTMRATSTMSGCAQTMIEHNSMTLES 360
Db 301 RHDEOORELEEEENEDEDELDSHTMVKTSVEGCTMRATSTMSGCAQTMIEHNSMTLES 360
Qy 361 DLGTWVINSDEEEEDGTYKRNATSPQVORPSFMDYFDKODPKNKSHENCONMHEPPM 420
Db 361 DLGTWVINSDEEEEDGTYKRNATSPQVORPSFMDYFDKODPKNKSHENCONMHEPPM 420
Qy 421 SKVFPDNNKRVQDGFDFLKNLSLEELQMRKALDPMMERIEELRORYTAKRROPILDA 480
Db 421 SKVFPDNNKRVQDGFDFLKNLSLEELQMRKALDPMMERIEELRORYTAKRROPILDA 480
Qy 481 MDAKKRRQONF 491
Db 481 MDAKKRRQONF 491

```

RESULT 2

```

ID 09J110 PRELIMINARY: PRT: 497 AA.
AC 09J110;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE STE20-LIKE KINASE MST2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RA Lee K.-K., Ohyama T., Yonehara S.;
RT MST2, a physiological caspase substrate, highly sensitizes apoptosis
RT both upstream and downstream of caspase activation."
RT Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AE271361; AAF5790.1;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR InterPro: IPR001245; Tyr_kin.
DR Pfam: PF000069; pkinase.1
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00220; S_TKc.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
KW ATP-binding; kinase; transferase.
SQ SEQUENCE 497 AA: 56855 MW: 9CDD365437581665 CRC64:

```

Query Match 95.2%; Score 2442; DB 11; Length 497;
 Best Local Similarity 95.0%; Pred. No. 1.3e-158;
 Matches 472; Conservative 8; Mismatches 11; Indels 6; Gaps 1;

```

Qy 1 MEOPPAKSKLKLSLSDSLTKOPEEVFDVLEKLGESYGVFAIKHESQVVAIKQVPV 60
Db 1 MEOPPAKSKLKLSLSDSLTKOPEEVFDVLEKLGESYGVFAIKHESQVVAIKQVPV 60
Qy 61 ESDLOELIKEISIMQOCDSPYVYKYSYFNKNTDLMIVMEYCGAGSVSDIIRLNKLTLE 120
Db 61 ESDLOELIKEISIMQOCDSPYVYKYSYFNKNTDLMIVMEYCGAGSVSDIIRLNKLTLE 120
Qy 121 DELATILKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGNAKLADPGVAGQLTDTMAKRN 180
Db 121 DELATILKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGNAKLADPGVAGQLTDTMAKRN 180

```

```

Db 121 DEATILKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGNAKLADPGVAGQLTDTMAKRN 180
Qy 181 VIGTFPMMAPEVIOEIGYNCVADIMSLGITSIEMAEKGPYPADIPHMAIEMIPNPPPT 240
Db 181 VIGTFPMMAPEVIOEIGYNCVADIMSLGITSIEMAEKGPYPADIPHMAIEMIPNPPPT 240
Qy 241 FRKPELMSDDFTDFPKKCLVKNPEORATATOLLHPFIKNKPPVSLRLDLTEAMEIKAK 300
Db 241 FRKPELMSDDFTDFPKKCLVKNPEORATATOLLHPFIKNKPPVSLRLDLTEAMEIKAK 300
Qy 301 RHDEOORELEEEENEDEDELDSHTMVKTSVEGCTMRATSTMSGCAQTMIEHNSMTLES 360
Db 301 RHDEOORELEEEENEDEDELDSHTMVKTSVEGCTMRATSTMSGCAQTMIEHNSMTLES 360
Qy 361 DLGTWVINS-----EDEEEDGTYKRNATSPQVORPSFMDYFDKODPKNKSHENCONM 414
Db 361 DLGTWVINSDEEEEDGTYKRNATSPQVORPSFMDYFDKODPKNKSHENCONM 420
Qy 415 HEPFMSKNVFPDNNKRVQDGFDFLKNLSLEELQMRKALDPMMERIEELRORYTAKR 474
Db 415 HEPFMSKNVFPDNNKRVQDGFDFLKNLSLEELQMRKALDPMMERIEELRORYTAKR 480
Qy 475 QPILDAMDAKKRRQONF 491
Db 475 QPILDAMDAKKRRQONF 491

```

RESULT 3

```

ID 060877 PRELIMINARY: PRT: 445 AA.
AC 060877;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PROTEIN KINASE MESS1.
GN STK3 OR MESS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Han J.;
RT Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL: U28726; AA875300.1;
DR HSSP: P00518; LPHK.
DR MGD: MGI:1928487; Stk3.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF000069; pkinase.1.
DR SMART: SM00220; S_TKc.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
KW ATP-binding; transferase.
SQ SEQUENCE 445 AA: 50407 MW: 7B936B28DA616F05 CRC64:

```

Query Match 82.3%; Score 2110; DB 11; Length 445;
 Best Local Similarity 94.9%; Pred. No. 4.9e-136;
 Matches 409; Conservative 6; Mismatches 10; Indels 6; Gaps 1;

```

Qy 1 MEOPPAKSKLKLSLSDSLTKOPEEVFDVLEKLGESYGVFAIKHESQVVAIKQVPV 60
Db 1 MEOPPAKSKLKLSLSDSLTKOPEEVFDVLEKLGESYGVFAIKHESQVVAIKQVPV 60
Qy 61 ESDLOELIKEISIMQOCDSPYVYKYSYFNKNTDLMIVMEYCGAGSVSDIIRLNKLTLE 120
Db 61 ESDLOELIKEISIMQOCDSPYVYKYSYFNKNTDLMIVMEYCGAGSVSDIIRLNKLTLE 120
Qy 121 DELATILKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGNAKLADPGVAGQLTDTMAKRN 180
Db 121 DELATILKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGNAKLADPGVAGQLTDTMAKRN 180

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:03:42 ; Search time 26.93 seconds
(without alignments)
410.290 Million cell updates/sec

Title: US-09-822-110-2
Perfect score: 2564
Sequence: 1 MEQPPAPKSLKLUKLESDSLT.....AKROPILDAMDKKRRQCNF 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 212252 seqs, 22503292 residues

al number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1960.5	76.5	487	2	US-08-712-709-8
2	1960.5	76.5	487	3	US-09-111-444-8
3	1960.5	76.5	487	4	US-09-541-228-8
4	1333	52.0	270	2	US-08-852-743-5
5	1333	52.0	270	3	US-09-185-370-5
6	817	31.9	431	2	US-09-211-930-5
7	817	31.9	431	3	US-09-340-993-5
8	817	31.9	431	4	US-09-152-406-3
9	817	31.9	431	4	US-09-468-442-5
10	815.5	31.8	416	2	US-09-211-930-3
11	815.5	31.8	416	3	US-09-340-993-3
12	815.5	31.8	416	3	US-09-111-444-3
13	815.5	31.8	416	3	US-09-468-442-3
14	815.5	31.8	416	4	US-09-185-370-5
15	815.5	31.8	416	4	US-09-468-442-3
16	793.5	30.9	426	2	US-08-852-743-2
17	793.5	30.9	426	2	US-09-211-930-4
18	793.5	30.9	426	3	US-09-340-993-4
19	793.5	30.9	426	3	US-09-185-370-2
20	793.5	30.9	426	4	US-09-468-442-4
21	793.5	30.9	426	4	US-09-152-406-4
22	739	28.8	403	2	US-08-712-709-3
23	739	28.8	403	3	US-09-111-444-3
24	739	28.8	403	4	US-09-541-228-3
25	701.5	27.4	1360	4	US-09-393-569-2
26	645	25.2	1001	4	US-09-060-410-2
27	631	24.6	465	2	US-08-114-555A-2

28	629	24.5	545	2	US-08-935-760-4	Sequence 4, App11
29	628.5	24.5	524	2	US-08-615-942A-2	Sequence 2, App11
30	628	24.5	544	3	US-08-935-760-2	Sequence 2, App11
31	627	24.5	544	3	US-08-559-397A-19	Sequence 19, App1
32	626	24.4	276	2	US-08-852-743-7	Sequence 7, App11
33	626	24.4	276	3	US-09-185-370-7	Sequence 7, App11
34	620.5	24.2	506	1	US-08-369-780-2	Sequence 2, App11
35	620.5	24.2	506	1	US-08-475-682-2	Sequence 2, App11
36	620.5	24.2	506	1	US-08-780-833-2	Sequence 2, App11
37	620.5	24.2	506	1	US-08-636-036-2	Sequence 2, App11
38	620.5	24.2	506	3	US-08-918-509-2	Sequence 2, App11
39	620.5	24.2	506	3	US-09-108-262-2	Sequence 2, App11
40	619	24.1	465	3	US-08-559-397A-2	Sequence 2, App11
41	610	23.8	268	2	US-08-852-743-3	Sequence 3, App11
42	610	23.8	268	3	US-09-185-370-3	Sequence 3, App11
43	606	23.6	544	3	US-08-559-397A-29	Sequence 29, App1
44	601	23.4	544	3	US-08-559-397A-30	Sequence 30, App1
45	598	23.3	993	4	US-09-060-410-4	Sequence 4, App11

ALIGNMENTS

RESULT 1
US-08-712-709-8
Sequence 8, Application US/08712709
Patent No. 5863780
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER-READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTA Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,709
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1117791
US-08-712-709-8
Query Match 76.5% ; Score 1960.5 ; DB 2: Length 487;
Best Local Similarity 78.0% ; Pred. No. 1.3e-160;
Matches 379; Conservative 54; Mismatches 44; Indels 9; Gaps 5;
QY 7 PKSKLKLKLESDSLTKEVEFVLEKLGESGYSVFRAIKHESGQVVAIKQVPEVDLQ 66

```

Db 10 PRRLKRLDESLTKOPEEYFDVLEKLGEGSYGVAIHKETQIVAIKQVPESDLOE 69
QY 67 IIRKESIMQCCDSPYVVKYYSYFKNTDLMVMEYCGAGSVDILRLRNKTLFDEIATI 126
Db 70 IIRKESIMQCCDSPHYVVKYYSYFKNTDLMVMEYCGAGSVDILRLRNKTLFDEIATI 129
QY 127 LKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGHAKLADFGVAGQLTDTMAKRNIVIGTFP 186
Db 130 LQSTLKGLEYLHFMKRIHRDIKAGNILLNTEGHAKLADFGVAGQLTDTMAKRNIVIGTFP 189
QY 187 WMAPEVIOEIGYNCVADWISGITSIEMAEGRPPYADHHPKRALFMTPTNPPFRKPEL 246
Db 190 WMAPEVIOEIGYNCVADWISGITAIEMAEGRPPYADHHPKRALFMTPTNPPFRKPEL 249
QY 247 WSDPTDFVKKCLVKNPEORATATQLOHPFIKNKAPVSIIRDLTTEAMEIKARHDEQ 306
Db 250 WSDNTDFVKKCLVKNPEORATATQLOHPFVRSKAGVSIIRDLTTEAMEIKARHDEQ 309
QY 307 REL-EEBENSDEDELDSHTWKTSGVCGCTMRATSTSEAGQTMIEHNSITMESDICTM 365
Db 310 REMDQDDENSEDEMDSGTVRAVGDEMGTVRAVSTMTDANTMI EHDIT- LPSQLGTM 368
QY 366 VINSEDEEEDGCTMKRNATSPVOYRSPMDYFDKODFKNKSHECNQNMHEPFPMSKNVF 425
Db 369 VINADEDEEE-CTMKRRETIQPAKPSFLYEFO--KEKEMOINSFKSVGPKLNS-- 422
QY 426 PDNMKVPDGDGDFLKNLSLEELQRLKALDPMMERIEIEELRORYTAKROPILDAMDAKK 485
Db 423 -SDWKIPDGDGYEFLKMTVEDLQRLALDPMMEQIEIEIROKYOSKROPILDAIEAKK 481
QY 486 RROQNF 491
Db 482 RROQNF 487

```

RESULT 2
US-09-111-444-8
Sequence 8, Application US/09111444
Patent No. 6045792
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,444
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
INFORMATION CHARACTERISTICS:

```

LENGTH: 487 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1117791
US-09-111-444-8

```

Query Match 76.5%; Score 1960.5; DB 3; Length 487;
Best Local Similarity 78.0%; Pred. No. 1,3e-160;
Matches 379; Conservative 54; Mismatches 44; Indels 9; Gaps 5;

```

Db 10 PRRLKRLDESLTKOPEEYFDVLEKLGEGSYGVAIHKETQIVAIKQVPESDLOE 69
QY 67 IIRKESIMQCCDSPYVVKYYSYFKNTDLMVMEYCGAGSVDILRLRNKTLFDEIATI 126
Db 70 IIRKESIMQCCDSPHYVVKYYSYFKNTDLMVMEYCGAGSVDILRLRNKTLFDEIATI 129
QY 127 LKSTLKGLEYLHFMKRIHRDIKAGNILLNTEGHAKLADFGVAGQLTDTMAKRNIVIGTFP 186
Db 130 LQSTLKGLEYLHFMKRIHRDIKAGNILLNTEGHAKLADFGVAGQLTDTMAKRNIVIGTFP 189
QY 187 WMAPEVIOEIGYNCVADWISGITSIEMAEGRPPYADHHPKRALFMTPTNPPFRKPEL 246
Db 190 WMAPEVIOEIGYNCVADWISGITAIEMAEGRPPYADHHPKRALFMTPTNPPFRKPEL 249
QY 247 WSDPTDFVKKCLVKNPEORATATQLOHPFIKNKAPVSIIRDLTTEAMEIKARHDEQ 306
Db 250 WSDNTDFVKKCLVKNPEORATATQLOHPFVRSKAGVSIIRDLTTEAMEIKARHDEQ 309
QY 307 REL-EEBENSDEDELDSHTWKTSGVCGCTMRATSTSEAGQTMIEHNSITMESDICTM 365
Db 310 REMDQDDENSEDEMDSGTVRAVGDEMGTVRAVSTMTDANTMI EHDIT- LPSQLGTM 368
QY 366 VINSEDEEEDGCTMKRNATSPVOYRSPMDYFDKODFKNKSHECNQNMHEPFPMSKNVF 425
Db 369 VINADEDEEE-CTMKRRETIQPAKPSFLYEFO--KEKEMOINSFKSVGPKLNS-- 422
QY 426 PDNMKVPDGDGDFLKNLSLEELQRLKALDPMMERIEIEELRORYTAKROPILDAMDAKK 485
Db 423 -SDWKIPDGDGYEFLKMTVEDLQRLALDPMMEQIEIEIROKYOSKROPILDAIEAKK 481
QY 486 RROQNF 491
Db 482 RROQNF 487

```

RESULT 3
US-09-541-228-8
Sequence 8, Application US/09541228
Patent No. 6232077
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:04:27 ; Search time 32.41 Seconds

(without alignments)
37.605 Million cell updates/sec

Title: US-09-822-110-3

Perfect score: 81

Sequence: 1 EIKAKRHDEQORELEE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-68:*

1: pirl:1:
2: pirl:2:
3: pirl:3:
4: pirl:4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	59.3	799	2 T00331	hypothetical prote
2	47	58.0	2357	2 A59249	class VII unconven
3	46.5	57.4	222	2 T14674	hypothetical prote
4	46.5	57.4	233	2 T14940	hypothetical prote
5	45	55.6	329	2 T40419	mlf2 protein - fls
6	45	55.6	1392	2 A4336	microtubule-vesicl
7	45	55.6	1427	2 S23695	restin - human
8	44.5	54.9	455	2 S65157	hypothetical prote
9	44	54.3	1537	2 JC4172	DNA (cytosine-5'-)
10	44	54.3	1621	2 T15264	hypothetical prote
11	43	53.1	444	2 C75428	probable permease
12	43	53.1	499	2 D85433	globulin-like prot
13	43	53.1	522	2 C96608	hypothetical prote
14	43	53.1	528	2 S57835	19g-binding protei
15	43	53.1	643	1 T07064	seed blotin-contai
16	43	53.1	1219	2 T33927	hypothetical prote
17	43	53.1	1655	2 T32633	hypothetical prote
18	43	53.1	3225	2 I52300	giantin - human
19	43	53.1	3259	1 A56539	giantin - human
20	42	51.9	197	1 A27784	dermorphin precurs
21	42	51.9	198	2 B27784	dermorphin precurs
22	42	51.9	419	2 T40014	probable vesicular
23	42	51.9	488	2 T16402	hypothetical prote
24	42	51.9	890	1 F1RC2	translation initia
25	42	51.9	890	2 D85980	protein chain init
26	42	51.9	1069	2 T00377	KIAA0642 protein -
27	42	51.9	1195	1 C64440	DNA repair protein
28	42	51.9	1240	2 S52734	hypothetical prote
29	42	51.9	2152	2 T45583	hypothetical prote

30	41.5	51.2	458	2 S73658	MG096 homolog p02-
31	41.5	51.2	592	2 T46657	heat shock protein
32	41.5	51.2	596	2 C73385	dnak-type molecule
33	41	50.6	143	2 A71717	hypothetical prote
34	41	50.6	180	1 GCHU	glucagon precursor
35	41	50.6	180	1 GCCP	glucagon precursor
36	41	50.6	180	1 GCRTDU	glucagon precursor
37	41	50.6	180	1 GCRTU	glucagon precursor
38	41	50.6	180	1 GCHY	glucagon precursor
39	41	50.6	180	1 GCBO	glucagon precursor
40	41	50.6	180	2 A57294	glucagon precursor
41	41	50.6	478	2 S51457	probable membrane
42	41	50.6	527	2 F64210	hypothetical. prote
43	41	50.6	602	2 H82841	conserved hypothet
44	41	50.6	699	2 E84565	hypothetical prote
45	41	50.6	1355	2 T32092	hypothetical prote

ALIGNMENTS

RESULT 1
T00331
hypothetical protein KIAA0555 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00331
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, DNA Res. 5, 31-39, 1998
A>Title: Prediction of the coding sequences of unidentified human genes. IX. The comp
A:Reference number: Z14086; MUID:98290545
A:Accession: T00331
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-799 <NAC>
A:Cross-References: EMBL:AB011127; NID:93043633; PIDN:BA25481.1; PID:93043634
A:Experimental source: brain; clone HH0882
C:Genetics:
A>Note: KIAA0555

Query Match 59.3%; Score 48; DB 2; Length 799;
Best Local Similarity 53.3%; Pred. No. 17;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EIKAKRHDEQORELE 15
DB 72 EIKAKRHDEQORELE 86

RESULT 2
A59249
class VII unconventional myosin - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
C:Accession: A59249
R:Titus, M.A.; Kuspa, A.; Loomis, W.F.
Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994
A>Title: Discovery of myosin genes by physical mapping in Dictyostelium.
A:Reference number: Z20873; MUID:95023928
A:Accession: A59249
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-2357 <TTT>
A:Cross-References: GB:L35321; NID:96226760; PIDN:AAF06035.1; PID:96226761
A:Experimental source: gene myoI; product class VII unconventional myosin
R:Titus, M.A.
Curr. Biol. 9, 1297-1303, 1999
A>Title: A class VII unconventional myosin is required for phagocytosis.
A:Reference number: A59249
A:Contents: annotation
C:Genetics:
A:Gene: myoI

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:03:42 ; Search time 54.15 Seconds

(without alignments)
671.653 Million cell updates/sec

Title: US-09-822-110-2

Perfect score: 2364
Sequence: 1 MEQPPAPSKLKKLSLSESLT.....AKROPILDMADAKRRQNF 491

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

al number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1957.5	76.3	487	AAV21674	Human Ste20 homolo
2	817	31.9	431	AAV55950	Human MEF3 protein
3	817	31.9	431	AAV04473	Human STE20-like s
4	817	31.9	431	AAV82276	Human STE20-like k
5	817	31.9	431	AAV68771	Amino acid sequenc
6	817	31.9	431	AAV87069	Human polypeptide
7	815.5	31.8	416	AAV55927	Human STRK2 protei
8	815.5	31.8	416	AAV22648	A human signal tra
9	815.5	31.8	416	AAV22651	A murine signal tr
10	815.5	31.8	416	AAV21673	Murine 5e.new poly
11	815.5	31.8	416	AAV82274	Human Ste20-like s

12	815.5	31.8	416	21	AAV82277	Mouse protein sequ
13	815.5	31.8	416	22	AAV40348	Human polypeptide
14	815.5	31.8	416	22	AAV71958	Human TGF-beta rec
15	815.5	31.8	416	22	AAV66607	Human h2252 protei
16	814.5	31.8	413	20	AAV21672	Murine C12.2bs pol
17	800	31.2	443	20	AAV55951	Nematode STE20-rel
18	796	31.0	426	20	AAV55949	Human STE20 protei
19	793.5	30.9	426	21	AAV82275	Human MAPK-pathway
20	789	30.8	426	18	AAV21603	Human protein kin
21	739	28.8	403	19	AAV54004	Human protein kin
22	714.5	27.9	1202	21	AAV808521	A murine Ste20-rel
23	705.5	27.5	1615	22	AAV03536	Human protein kin
24	701.5	27.4	1298	22	AAV86822	Amino acid sequenc
25	701.5	27.4	1306	22	AAV868218	Amino acid sequenc
26	701.5	27.4	1353	22	AAV868219	Amino acid sequenc
27	701.5	27.4	1360	21	AAV85263	Human protein kin
28	701.5	27.4	1360	22	AAV80059	Large NIK-related
29	698	27.2	1269	22	AAV868223	Amino acid sequenc
30	698	27.2	1277	22	AAV868220	Amino acid sequenc
31	698	27.2	1324	22	AAV86821	Amino acid sequenc
32	698	27.2	1332	22	AAV868217	Amino acid sequenc
33	697.5	27.2	912	20	AAV55939	Human CERK2 protein
34	697.5	27.2	968	20	AAV55966	Full length human
35	693	27.0	911	20	AAV55957	Mouse STE20-relate
36	693	27.0	1212	22	AAV5931	Human kinase (PKIN
37	693	27.0	1239	20	AAV59316	Human ZC1 protein.
38	692	27.0	1165	21	AAV43016	Human OREF ORF2780
39	691	27.0	1135	21	AAV68784	Amino acid sequenc
40	678.5	26.5	894	20	AAV55935	Human KHS2 protein
41	666	26.0	842	20	AAV55956	Human STE20-relate
42	664.5	25.9	1233	20	AAV55954	Mouse STE20-relate
43	660	25.7	1106	22	AAV03552	Human protein kin
44	649	25.3	1297	20	AAV55932	Human ZC2 protein.
45	645	25.2	1001	20	AAV49896	Rat TA01 kinase.

ALIGNMENTS

RESULT 1
AAV21674
ID AAV21674 standard; Protein: 487 AA.
XX
AC AAV21674;
XX
XX 18-AUG-1999 (first entry)
DT
XX Human Ste20 homologue polypeptide S201.
DE
XX Sterile 20; Ste20; serine/threonine kinase; C12.2bs; 5e.new; murine;
KW erythroleukemia; mitogenic signaling; cell differentiation; metastasis;
KW signaling pathway; cancer; ischemic stroke; heart disease; inflammation.
XX
XX Homo sapiens.
OS
XX
XX MO9929857-A1.
XX
XX 17-JUN-1999
XX
XX 09-DEC-1998; 98WO-US26116.
XX
XX 09-DEC-1998; 98WO-US26116.
XX
XX 09-DEC-1997; 97US-0069078.
XX
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX
XX Agarwal S, Best J, Vail B, Zon LI;
XX WPI; 1999-385605/32.
XX
XX Murine homologues of the Sterile 20 family of serine/threonine
XX kinases, useful in elucidation of mitogenic signaling pathways
XX
XX Disclosure; Fig 4; 37pp; English.
PS

Query Match	76.3%;	Score 1957.5;	DB 20;	Length 48/;
Best Local Similarity	77.2%;	Pred. No. 2.2e-161;		
Matches 379;	Conservative	54;	Mismatches 39;	Indels 19;
				Gaps 5

XX	antiproliferative; antiarteriosclerotic; antiasthmatic; immunosuppressive;
KW	neuroprotective; cardiant; cerebroprotective; cytostatic; antiabiotic;
KW	vincetristine; protein kinase; STLK2; STLK3; STLK4; STLK5; STLK6; STLK7;
KW	ZC1, ZC2, ZC4, KHS2, SUI01, SUI03, GEXK, PAK1, PAK5; antagonist;
KW	antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;
KW	inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
KW	rhinitis; autoimmunity; organ transplantation; multiple sclerosis;
KW	myocardial infarction; cardiovascular disease; stroke; renal failure;
KW	oxidative stress-related neurodegenerative disorder; Parkinson's disease;
KW	amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
KW	ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;
KW	mesangial disorder; growth regulation; diabetes mellitus; T cell activation;
KW	immunosuppressant.
OS	Homo sapiens.
XX	WO9953036-A2.
PN	21-OCT-1999.
XX	13-APR-1999; 99WO-US08150.
PF	14-APR-1998; 98US-0081784.
PR	(SUGEN-) SUGEN INC.
XX	Plowman G, Martinez R, Whyte D;
PI	WPI, 1999-611301/52.
DR	Novel kinase-related polypeptides used for the diagnosis and treatment
XX	of kinase-related diseases and disorders
PT	Disclosure; Page 330-332; 387pp; English.
XX	
PS	
CC	This sequence represents a novel STK20-related protein kinase. The
CC	invention relates to nucleic acid molecule encoding a kinase polypeptide
CC	selected from STLK2, STLK3, STLK4, STLK5, STLK6, STLK7, ZC1, ZC2, ZC3,
CC	ZC4, KHS2, SUI01, SUI03, GEXK2, PAK4 and PAK5. The proteins are used to
CC	identify agonists and antagonists, and to raise antibodies. The
CC	polynucleotides are useful in gene therapy protocols. The polynucleotides,
CC	polynucleotides, antibodies, antagonists and agonists may be used to treat
CC	diseases such as immune-related disorders and diseases (e.g., rheumatoid
CC	arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g.,
CC	Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,
CC	rhinitis, autoimmunity, and organ transplantation, chronic inflammatory
CC	pelvic disease, multiple sclerosis, organ transplantation, myocardial
CC	infarction, cardiovascular disease, stroke, renal failure, oxidative
CC	stress-related neurodegenerative disorders (e.g., amyotrophic lateral
CC	sclerosis, Parkinson's disease and Leigh syndrome), cancer, diabetes
CC	cardiomyopathy, ischemic disorders, inflammatory disorders, diabetes
CC	mellitus, fibrotic and mesangial disorders. The proteins may also be
CC	useful for cell growth regulation (e.g., in wound healing), T cell
CC	activation, mitosis control, and as immunosuppressants.
XX	
XX	Sequence 431 AA;
XX	
QY	Query Match 31.9%; Score 817; DB 20; Length 431;
DB	Best local Similarity 39.3%; Pred. No. 1.6e-62;
DB	Matches 188; Conservative 73; Mismatches 123; Indels 94; Gaps 13;
QY	18 SLTKPQEFVDELKLGSGSYGSKAIIHKSQGVAKIKQPV--ESDLOEIIKEISIM 74
DB	15 nlkadpeelftkxkysgsvfgrkgjldntqkvaklidldeaeadedidqgetl 74
QY	75 QQCDDSPYVAVKSYFKNTDMLIVWYEGAGSVSDIIRLNKTLIDELITFIKSTFKG 134
DB	75 sgcdspvyvkkygyslkdtklwlmeysl999saldl--leggldedqalallrelkgl 132
QY	135 EYLHMRKTHRDKACGNILNTEGNAKLADGCVAGOLDTMAKNFTVIGTFPMWAPVIO 194
DB	133 dylhsekhlrdkaavlllsegevkladiqvagldcqlkntftvgpfrfmwapevik 192

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2002, 07:03:42 ; Search time 32.41 Seconds

(without alignments)
1154.017 Million cell updates/sec

Title: US-09-822-110-2

Perfect score: 2566

Sequence: 1 MEOPAPKSKLKLSSESLT.....AKROPILDAMAKRRQNF 491

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	800	31.2	653	2 T34356	hypothetical prote
2	793.5	30.9	426	2 S71886	Stre20-1like protei
3	774.5	30.2	836	2 B96716	probable serine/th
4	714.5	27.9	1233	2 T14157	serine/threonine p
5	701.5	27.4	1231	2 T18532	serine/threonine p
6	697	27.2	471	2 T39232	probable serine th
7	696	27.1	1206	2 T34021	protein kinase SK2
8	682	26.6	690	2 C96572	protein kinase SK2
9	664.5	25.9	1233	2 T30989	serine/threonine p
10	657	25.6	1102	2 JC6316	probable protein k
11	648.5	25.3	829	2 T29372	hypothetical prote
12	646	25.2	819	2 A53714	protein kinase (EC
13	645	25.2	1001	2 T17365	serine/threonine p
14	634	24.7	544	2 S40482	serine/threonine-s
15	633	24.7	545	2 G01773	p21-activated prot
16	630	24.6	544	2 A57597	beta-p21-activated
17	629	24.5	544	2 T49376	p21-activated kina
18	628	24.5	544	2 T38525	serine/threonine p
19	627	24.5	525	2 S58682	protein kinase, p2
20	618.5	24.1	982	2 T18576	serine/threonine k
21	595.5	23.2	658	2 T39500	serine/threonine-s
22	593	23.1	1075	2 T27623	hypothetical prote
23	592	23.1	1080	2 T27622	hypothetical prote
24	593	23.1	490	2 S47946	protein kinase hom
25	566.5	22.1	658	2 S60170	protein kinase pak
26	566.5	22.1	1228	2 T18897	hypothetical prote
27	565.5	22.1	652	2 T39722	serine/threonine p
28	564.5	22.0	378	2 T26684	hypothetical prote
29	564.5	22.0	939	2 S28394	probable serine/th

30	562	21.9	632	2 T15467	hypothetical prote
31	554	21.6	1230	2 T18256	probable serine/th
32	554	21.6	1230	2 T18259	serine/threonine p
33	552	21.5	589	2 T38086	serine/threonine-p
34	552	21.5	1080	2 S48944	hypothetical prote
35	549.5	21.4	1135	1 A29813	132K ninac protein
36	549.5	21.4	1501	1 B29813	174K ninac protein
37	547	21.3	533	2 T01479	hypothetical prote
38	540	21.1	561	2 T51417	protein kinase-11k
39	538	21.0	693	2 B85112	hypothetical prote
40	531	20.7	108	2 T38212	protein-serine/thr
41	526.5	20.5	842	2 S60402	protein kinase CIA
42	520	20.3	607	2 T01904	hypothetical prote
43	518.5	20.2	1014	2 T31109	myosin III - Aclan
44	516.5	20.1	1062	2 S46367	protein kinase CDC
45	509.5	19.9	655	2 S51884	probable protein k

ALIGNMENTS

RESULT	1	Score	DB 2	Length	653
T34356	hypothetical protein T19A5.2 - Caenorhabditis elegans	31.2%			
C:Species: Caenorhabditis elegans		Best Local Similarity 14.3%: Pred. No. 3.1e-24;			
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999		Matches 163; Conservative 68; Mismatches 102; Indels 62; Gaps 6;			
C:Accession: T34356					
R:Bradshaw, H.					
Submitted to the EMBL Data Library, March 1996					
A:Description: The sequence of C. elegans cosmid T19A5.					
A:Reference number: 221512					
A:Accession: T34356					
A:Molecule type: DNA					
A>Status: preliminary; translated from GB/EMBL/DBJ					
A:Residues: 1-653 <BBA>					
A:Cross-references: EMBL:U53153; PDB:1AC69038.1; GSPDB:GM00023; CESP:T19A5.2					
A:Experimental source: strain Bristol NZ; clone T19A5					
C:Genetics:					
A:Gene: CESP:T19A5.2					
A:Map position: 5					
A:Introns: 26/3; 101/3; 157/1; 209/3; 272/3; 305/2; 373/3; 635/3					
Query Match					
Best Local Similarity 14.3%: Pred. No. 3.1e-24;					
Matches 163; Conservative 68; Mismatches 102; Indels 62; Gaps 6;					
OY	15 SEDSLTKQ-----	-----	PEEVFDVLEKLGSGSVFKAIHKESGOVAIKQ	57	
DB	5 SSDLPRQADDSSMKWDRIYIQKLDPEVIFPKOERIGSGFGEYVKGIDNTRGVVAIKI	64			
OY	58 VPV---ESDLOEIKETISIMGDCSPYVVKYKSGFTMDIMVWEXGAGSVSIILRL	114			
DB	65 IDLEQAEDEIDIDOEIVLQSCDSQYVTKTFGSLKSKIMIMETIYIGGSALDL--TK	122			
OY	115 NKTLEDEIATILKSTKLEYLEHMRKIRHDINAGNILLTEGHAKLADEGVAGQLTDT	174			
DB	123 SGKIDESHIAVILREILKLEYLEHMRKIRHDINAGNILLTEGHAKLADEGVAGQLTDT	182			
OY	175 MAKRTVIGTFPMAPPEVIOETGNCVADIMSLGITSLEMAEGKPPYADHPRNAIFIP	234			
DB	183 VKKRTTFVSGPFMAPELIKOSYDKADIMSLGITAELANGEPHSDHPRNAIFIP	242			
OY	235 TNPPTTFKPELMSDDFTDFVKCKLVKPEORATATOLQHPFIKNAPVSIILRLDTEA	294			
DB	243 KNPPEVLQSGQ--WSKPFKEFVEMCLNKDPENRPSATILKHQFIKRAKNSILVDLIERA	301			
OY	295 MEIKAKRHDEQORELEEEENSDDELDSHTVMTVTSVGECSGTMRATSTMSGAQTMIEHN	354			
DB	302 AEYRLRTGVSSDSDDLDESDGG-----GTSKMDYPTVVGPRVSA-----	341			
OY	355 STMLESIDLGTVINSEDEEDGCTMKRNATSPQVQ 389				

Db 342 -----DDGGVVRORTDPRRAQ 357

RESULT 2

Ste20-like protein kinase - human

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 21-Jul-2000

C:Accession: S71886; G01022

R:PubMed, C.M.; Bonventre, J.V.; Molnar, A.; Kyriakis, J.; Force, T.

EMBO J. 15, 4537-4546, 1996

A:Title: Activation of a human Ste20-like kinase by oxidant stress defines a novel stress

A:Reference number: S71886; M01D:97042345

A:Accession: S71886

A:Molecule type: mRNA

A:Residues: 1-426 <POM>

A:Cross-references: EMBL:X99325; NID:q1430821; PIDN:CAA67700.1; PID:q1430822

A:Experimental source: cell type B cell

A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Keywords: ATP; autophosphorylation; phosphotransferase; protein kinase; stress-induced

F:18-270/domain: protein kinase homology <kin>

Query Match 30.9%; Score 793.5; DB 2; Length 426;

Best Local Similarity 52.2%; Pred. No. 3.8e-24;

Matches 157; Conservative 53; Mismatches 78; Indels 13; Gaps 5;

23 PEEVFDVLEKIGEGSYGSGVSKAIHKSQGVVAIKOVV--ESDLOEIKETISIMQCCDS 79

16 PEEFTKLDRIKSGSGFEVYKGIIDNHKEVVAIKIIDELEAEDEIDIOQETVLVSQCCDS 75

80 PYVVKYIGSYRKNDIMTWMEYCGAGSVSDILRNKLTIDEIATILKSTLGLLEYLHF 139

76 PYTRKFEYSYKSKYKIMTYMEYLGSSALD--LKRGPLEEYIATILRELKGLDYLSH 133

140 MKRIHDKAGNINILNTEGHAKLADFVAGOLDFTMKRNVTIGTPPMAPVYIOELGYN 199

134 ERKIHDKANVILNTEGHAKLADFVAGOLDFTMKRNVTIGTPPMAPVYIOELGYN 193

200 CVADISIGTISIMAEKRPYADIHPRAFEMIPTPNPTFRPELMSDFDVEYKCL 259

194 FRADISIGTISIMAEKRPYADIHPRAFEMIPTPNPTFRPELMSDFDVEYKCL 251

260 VKNPEGRATQLOHPFI-KNAKPVSLRLDITAMEITAKRDEOORELEEEENSDE 318

252 NKDPRRPRAKELKHKFTTFTYKTSFLTELIDRYKMKSGHGCE-----ESSSESDI 306

319 D 319

307 D 307

RESULT 3

B96716

probable serine/threonine kinase F23010.20 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B96716

R:PubMed, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

R:PubMed, A.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.

Chin, C.W.; Hughes, B.; Hultner, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; M01D:21016719

A:Accession: B96716

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 836 <STO>

A:Cross-references: GB:AE005173; NID:q7705100; PIDN:AAF67779.1; GSPDB:GN00141

C:Genetics:

A:Gene: F23010.20

A:Map position: 1

Query Match 30.2%; Score 774.5; DB 2; Length 836;

Best Local Similarity 38.1%; Pred. No. 3.7e-23;

Matches 177; Conservative 73; Mismatches 143; Indels 71; Gaps 12;

4 PPAKSKL-----KLTSE-----DSLTQK-PEEVFDVLEKIGEGSYGSGVFAIH 46

209 PSSSSSKLHGENDNRKMQQNSMTSLPSDITREDPTTYKFEINELKSGSYGVSKARD 268

47 KESGQVVAIKOVV--ESDLOEIKETISIMQCCDSPVVKYSGYRKNDIMTWMEYCG 103

269 LKTSLEIVAKVSLTEGEGYEIEGELMLOQCHHPVAVYLSYCGEDVLWMEYCG 328

104 AGSVSDILRNKLTIDEIATILKSTLGLLEYLHFMRKIHDKAGNINILNTEGHAKLA 163

329 GGSVADLMNVTLEALFEYGLAICREALKGLAYLSIYKVRHDKGNILNTEGEGYKLG 388

164 DFGVAGOLDFTMKRNVTIGTPPMAPVYIOELGYNCAVDISIGTISIMAEKRPYAD 223

389 DFGVAGOLDFTMKRNVTIGTPPMAPVYIOELGYNCAVDISIGTISIMAEKRPYAD 448

224 IHPRAFEMIPTPNPTFRPELMSDFDVEYKCLVKNPEGRATQLOHPFIKNAKP 283

449 VHPKRVLEMTSIEPAPLEKESKSLVHFDFVAKCLTKERLRPTAAEMLKHKFVERCKT 508

284 -VSILRLDITAMEITAKRDEOORELEEEENSDEDELDSHTMVKTSVGECCG----- 335

509 GASMSKRIKESRQIRATMALQASVAPSL-----DTSILPKSSSEELGIVPSKP 561

336 -----TMRATSTMBEGAGTMEH--NSTMLE-----SDLGTVINISEDEEDSTMKRNAT 385

562 PNSTEADPTSTLNR-----QHTGTVLAGSGGDTGTVIGEDFETESDSRSOLVRE 615

386 PVOVRPFMDYFDKODKRNKSHENCONMHEPPMKNVYPDW 429

616 KESSSCF-----EGVPRFEGEE--LPDSM 639

RESULT 4

T14157

serine/threonine protein kinase - mouse

C:Species: Mus musculus (house mouse)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T14157

R:PubMed, B.; Hicklin, D.J.; Kornhaber, G.; Dellaratta, D.V.; Witte, L.

submitted to the EMBL Data Library, December 1997

A:Reference number: Z17894

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1233 <PVT>

A:Cross-references: EMBL:AF039574; NID:q2773155; PID:q2773156; PIDN:AB96682.1

Query Match 27.9%; Score 714.5; DB 2; Length 1233;

Best Local Similarity 32.0%; Pred. No. 1e-20;

Matches 171; Conservative 110; Mismatches 163; Indels 91; Gaps 14;

10 KLRKLSDSLTQK-----PEEVFDVLEKIGEGSYGSGVSKAIHKSQGVVAIKOVV 60

8 KIFELGSEKKKKYEHVKKDLNPEEWEIIGELGDGAPFKYKAKNKENVLAAKVIDT 67

61 -ESDLOEIKETISIMQCCDSPVVKYSGYRKNDIMTWMEYCGAGSVSDILRNKTL 118

68 KSEEDLWYVEIDLASCCHPNIVKLDFAFYENNMLWILIEFCAGGAVDAVWLELERL 127

119 IDEIATILKSTLGLLEYLHFMRKIHDKAGNINILNTEGHAKLADFVAGOLDFTMKAR 178